Australian Government
Department of Defence
Science and Technology

A Review of Sparsity-Based Methods for Analysing Radar Returns from Helicopter Rotor Blades

Ngoc Hung Nguyen[1], *Hai-Tan Tran*[2],
Kutluyıl Doğançay[1] *and Rocco Melino*[2]

[1] **University of South Australia**
[2] **National Security and ISR Division**
Defence Science and Technology Group

DST-Group–TR–3292

ABSTRACT

Radar imaging of rotating blade-like objects, such as helicopter rotors, using narrowband radar has lately been of significant interest; these objects cannot be adequately described by the classic point-scatterer model. Recently, a novel 'tilted-wire' scatterer model has been developed that can provide an accurate and sparse representation of radar returns from such objects.

Following a literature review on compressed sensing algorithms, covering both greedy and l_p minimisation methods ($0 < p \leq 1$), the report focuses on a comparative study of various greedy pursuit algorithms, using both simulated and real radar data, with a particular emphasis on the use of the tilted-wire scatterer model. It is observed that the greedy algorithms that select multiple atoms at the matched-filtering stage do not perform well when the atoms used in the dictionary are significantly correlated. Amongst the greedy algorithms, Orthogonal Matching Pursuit (OMP) exhibits the best performance, closely followed by Conjugate Gradient Pursuit (CGP), which has a much smaller computational complexity than OMP. In applications where the tilted-wire model requires large dictionaries and large CPI atoms, CGP is the preferred option.

RELEASE LIMITATION

Approved for Public Release

Published by

National Security and ISR Division
Defence Science and Technology Group
PO Box 1500
Edinburgh, South Australia 5111, Australia

Telephone: *1300 333 362*
Facsimile: *(08) 7389 6567*

AR-016-687
September, 2016

APPROVED FOR PUBLIC RELEASE

A Review of Sparsity-Based Methods for Analysing Radar Returns from Helicopter Rotor Blades

Executive Summary

Signal analysis and radar imaging of fast-rotating objects such as helicopter rotor blades are of particular research interest because these micro-Doppler signals cannot be processed by conventional range-Doppler techniques and should be separated from other non-rotating scattering components prior to further processing. In addition, the point-scatterer model, which is commonly assumed in the development of various inverse SAR (ISAR) and radar tomography approaches, is not always the most fitting model for the analysis of this type of micro-Doppler signals. The novel tilted-wire model offers new possibilities to overcome the limitations of the point-scatterer model; it however also introduces a new degree of complexity which requires the use of state-of-the-art sparsity-based techniques.

Since the tilted-wire scatterer model can be used to facilitate an accurate sparse representation of signals from rotating blades, a comprehensive review of known algorithms for sparse parameter estimation techniques is carried out, covering both greedy and l_p minimisation methods ($0 < p \leq 1$). The report focuses on a comparative study of various greedy pursuit algorithms using both simulated and real helicopter radar data, with an aim to accurately estimate the tilted-wire parameters associated with a rotor blade. These parameters are presented as scatter plots which show the orientation, length and tilt of the estimated wires used to represent the rotor blade.

Amongst the greedy algorithms, the so-called Orthogonal Matching Pursuit (OMP) technique exhibits the best performance, closely followed by Conjugate Gradient Pursuit (CGP), which has a much smaller computational complexity than OMP. Important improvements and exploitation of these modern techniques will be published separately in the near future.

THIS PAGE IS INTENTIONALLY BLANK

Contents

Figures

Tables

1 Introduction

Compressed sensing (CS) has recently received remarkable interest thanks to its wide range of applications in different research fields including electrical engineering, applied mathematics and computer science. In radar imaging, CS techniques have been widely exploited in many different problems such as sparse phase coherent imaging [1–3], wide-angle synthetic aperture radar (SAR) imaging for anisotropic scattering [4–7], multichannel SAR imaging [8, 9], and moving target indication [10, 11]. In this report, we are especially interested in the problem of CS-based radar imaging of rotating blade-like objects and the performance of different techniques for this problem.

Radar signals returned from rotating blade-like objects are usually grouped under the general category of 'micro-Doppler' signals, where the Doppler frequency modulation due to the blade rotation can be very complex and extensive [12–14]. Such micro-Doppler signals can not be processed by conventional range-Doppler techniques and should be separated from other non-rotating scattering components prior to further processing. In addition, the point-scatterer model, which is commonly assumed in the development of different inverse SAR (ISAR) and radar tomography approaches, are not always the most fitting model for the analysis of micro-Doppler signals.

In contrast to point-scatterers with isotropic scattering characteristics, rotating blade-like objects behave similarly to radial antennas with a lobe structure in their reflectivity pattern. All conventional radar imaging theories based on the ideal point-scatterer are thus not directly applicable. A plausible extension to the point-scatterer model is the 'tilted-wire scatterer' model, where the basic scatterer is modeled as a uniform straight wire characterised by the position of its centre and its shape [13, 14]. The centre position vector may be in 2D or 3D space, while the shape parameters include at least the finite length and the tilt angle of the wire relative to the radial direction.

Rotating blades, in particular, can be approximated as a collection of such scatterers, motivated by the fact that the blades may exhibit complex bending and twisting during high-speed rotational motion, and thus a strictly radial wire model may not be sufficient to characterise the scattering from the blades. More importantly, the tilted-wire model facilitates the application of CS techniques to the problem of radar imaging of rotating blades [14].

The fundamental problem in compressed sensing is to recover an unknown vector $\boldsymbol{x} \in \mathbb{C}^N$ from a small number of noisy linear measurements

$$\boldsymbol{y} = [y_1, \ldots, y_M]^T = \boldsymbol{\Phi}\,\boldsymbol{x} + \boldsymbol{n} \in \mathbb{C}^M, \quad (M \ll N). \tag{1}$$

Here, $\boldsymbol{\Phi} \in \mathbb{C}^{M\times N}$ is a known sensing matrix with columns normalised to unity and $\boldsymbol{n}$ is a noise vector with energy bound of $\|\boldsymbol{n}\|_2 < \epsilon$. Since the sensing matrix $\boldsymbol{\Phi}$ provides an over complete basis, a unique solution cannot be determined using the conventional inverse transform of $\boldsymbol{\Phi}$. However, if $\boldsymbol{x}$ is sparse or compressible (i.e., well-approximated as being sparse), CS theory enables the recovery of $\boldsymbol{x}$ from very few measurements in an effective and robust manner [15–24]. The sparse recovery problem is formulated as finding the sparsest solution of $\boldsymbol{x}$:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{C}^N}{\arg\min} \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad \boldsymbol{x} \in \mathcal{B}(\boldsymbol{y}) \tag{2}$$

where the constraint $\boldsymbol{x} \in \mathcal{B}(\boldsymbol{y})$ ensures that the solution $\hat{\boldsymbol{x}}$ is consistent with the measurements $\boldsymbol{y}$. If the measurements are noise-free, we can set

$$\mathcal{B}(\boldsymbol{y}) = \{\boldsymbol{x} : \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{y}\}. \tag{3}$$

Otherwise, if the measurements are contaminated by bounded noise, we can instead set

$$\mathcal{B}(\boldsymbol{y}) = \{\boldsymbol{x} : \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2 < \epsilon\}. \tag{4}$$

Solving the l_0 minimisation problem (2) is NP-hard, i.e., a highly non-convex combinatorial optimisation with exponential complexity, and thus is computationally intractable for practical applications [15–24]. An attractive alternative is to consider an l_1-minimisation problem

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{C}^N}{\arg\min} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{x} \in \mathcal{B}(\boldsymbol{y}), \tag{5}$$

which, in contrast to (2), is a convex optimisation problem (given that $\mathcal{B}(\boldsymbol{y})$ is convex) and can be solved effectively in polynomial time with standard convex optimisation techniques [20, 21, 23]. Importantly, it is well-known in the literature that the l_1-minimisation (5) is equivalent to the l_0-minimisation (2) for the sensing matrices satisfying the so-called restricted isometry property (RIP) with a constant parameter [16–19, 23, 24]. A matrix $\boldsymbol{\Phi}$ satisfies the RIP of order K if there exists a $\delta_K \in (0, 1)$ such that

$$(1 - \delta_K)\|\mathbf{x}\|_2^2 \leq \|\boldsymbol{\Phi}\mathbf{x}\|_2^2 \leq (1 + \delta_K)\|\mathbf{x}\|_2^2 \tag{6}$$

holds for all K-sparse vectors $\mathbf{x}$, i.e., vectors with at most K nonzero components.

Although the l_1-minimisation formulation based on convex optimisation provides a powerful framework for developing computationally tractable CS algorithms, the complexity of l_1-minimisation is still prohibitive for real-time applications [24–26]. Iterative greedy algorithms [23–32] have recently received significant attention as attractive alternatives to convex optimisation techniques thanks to their low complexity and simple geometric interpretation.

The main principle behind the greedy algorithms is that they identify the support of $\boldsymbol{x}$ iteratively, where one or more elements are selected at each iteration based on some greedy rules and their contribution is subtracted from the measurement $\boldsymbol{y}$. The main advantage of the greedy algorithms is that they are more computationally efficient than the l_1-minimisation algorithms. In terms of performance, some of the greedy algorithms have been shown to have theoretical performance guarantees that are comparable to those guarantees derived for convex l_1-norm optimisation approaches [21].

Although the CS literature has primarily focused on the l_1-minimisation and greedy pursuit algorithms, there exists another class of sparse recovery algorithms based on l_p-minimisation with $p < 1$. The l_p-minimisation with $p < 1$ was shown empirically in [33] to produce exact recovery with fewer measurements than with $p = 1$. It was also theoretically proven in terms of the RIP of $\boldsymbol{\Phi}$ that the sufficient condition for exact recovery of $\boldsymbol{x}$ via the l_p-minimisation is weaker for smaller p [34]. Although the l_p-minimisation with $p < 1$ is a non-convex problem and thus computationally intractable as described in the literature [34], the l_p-minimisation with $p < 1$ can be effectively solved via iteratively re-weighted least squares (IRLS) [34, 35].

In this report, we first review these three classes of sparse recovery algorithms based on greedy pursuit, l_1-minimisation, and l_p-minimisation with $p < 1$. We then focus our attention on the greedy pursuit algorithms and present a comparative performance study of these algorithms in the particular problem of analysing backscatter signals from rotating blades. The report is organised as follows. A survey on the CS literature is presented in Section 2. The comparative performance study among various greedy pursuit algorithms for radar imaging of rotating blades is presented in Section 3. The report concludes in Section 4 with a brief summary of the main findings.

2 Summary of Compressed Sensing Algorithms

This section summarize all three major classes of algorithms known as greedy, l_1-minimisation, and l_p-minimisation algorithms. However, only the greedy algorithms will be studied and discussed in great detail.

2.1 Greedy Algorithms

The greedy algorithms, which are conceptually simple and fairly straightforward, identify the support of the unknown sparse vector $\mathbf{x}$ progressively. These algorithms usually start with an initial estimate of

$$\hat{\mathbf{x}}^{[0]} = \mathbf{0}, \tag{7}$$

i.e. the support set – the indices of nonzero elements – of the initial estimate $\hat{\mathbf{x}}^{[0]}$ is $\Lambda = \emptyset$, and an initial residual of

$$\mathbf{r}^{[0]} = \mathbf{y} - \mathbf{\Phi}\,\hat{\mathbf{x}}^{[0]} = \mathbf{y}. \tag{8}$$

At each iteration, one or more columns of $\mathbf{\Phi}$ are selected based on the correlation values between the columns of $\mathbf{\Phi}$ and the residual $\mathbf{r}$. The indices of the selected columns are then added to the support set Λ, and the estimate $\hat{\mathbf{x}}$ as well as the residual $\mathbf{r}$ are also updated accordingly. The greedy algorithms repeat this procedure until a stopping criterion is triggered, discussed in more detail in Section 2.1.7.

A number of variants of iterative greedy algorithms have been published in the literature including the basic Matching Pursuit (MP) [27], Orthogonal Matching Pursuit (OMP) [28], Stagewise OMP (StOMP) [25], Stagewise Weak OMP (SWOMP) [29], Generalised OMP (gOMP) [26], Regularised OMP (ROMP) [23, 31], Compressive Sampling Matching Pursuit (CoSaMP) [32], Subspace Pursuit (SP) [24], Gradient Pursuit (GP) [30], Conjugate GP (CGP) [30], and Stagewise Weak CGP (SWCGP) [29]. Some of these variants will be discussed in further detail in the following sub-sections.

2.1.1 Basic Matching Pursuit (MP)

The simplest greedy algorithm is MP [27] which is summarised in Algorithm 1. At each iteration, MP selects one column of $\mathbf{\Phi}$, which yields the largest correlation with the current

Algorithm 1 MP Algorithm

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $j^* = \arg\max_j |\mathbf{c}_j|$
    Update:
      $\hat{\mathbf{x}}_{j^*}^{[i]} = \hat{\mathbf{x}}_{j^*}^{[i-1]} + \mathbf{c}_{j^*}$
      $\mathbf{r}^{[i]} = \mathbf{r}^{[i-1]} - \mathbf{\Phi}_{j^*} \mathbf{c}_{j^*}$
  end for
end procedure

residual, and updates the set of coefficients $\hat{\mathbf{x}}$ accordingly. Here, the coefficient update

$$\hat{\mathbf{x}}_{j^*}^{[i]} = \hat{\mathbf{x}}_{j^*}^{[i-1]} + \mathbf{c}_{j^*} \tag{9}$$

may occur for a coefficient that has been updated in an earlier iteration. In other words, MP may select a particular column of $\mathbf{\Phi}$ multiple times and refine its coefficient to enhance the approximation performance. The computational complexity of MP in each iteration is dominated by the matrix vector product $\mathbf{\Phi}^T \mathbf{r}^{[i-1]}$ in the identification step which requires a computation of $\mathcal{O}(MN)$ for unstructured matrices or $\mathcal{O}(N \log(M))$ for structured matrices by exploiting fast Fourier transform (FFT)-based algorithms [21, 30]. In addition, finding the largest element in $\mathbf{c}$ in the identification step and updating the residual $\mathbf{r}^{[i]}$ require N operations and M operations respectively.

2.1.2 Orthogonal Matching Pursuit (OMP)

The OMP algorithm [28], a more sophisticated refinement of MP, is one of the most popular greedy algorithms. The pseudo-code of OMP is summarised in Algorithm 2. The main difference between OMP and MP is that, in each iteration, OMP projects $\mathbf{y}$ orthogonally onto the columns of $\mathbf{\Phi}$ associated with the current support set $\Lambda^{[i]}$, in a least-squares-error sense, to obtain a new approximation of $\mathbf{x}$. In other words, OMP minimises $\|\mathbf{y} - \mathbf{\Phi}\hat{\mathbf{x}}^{[i]}\|_2^2$ over all $\hat{\mathbf{x}}^{[i]}$ with support $\Lambda^{[i]}$. Another difference is that OMP only selects an element maximally once and the residual $\mathbf{r}^{[i]}$ is always orthogonal to the current selected element set $\mathbf{\Phi}_{\Lambda^{[i]}}$ as a result of the least-squares estimation. This least-squares estimation makes OMP computationally more demanding than MP. However, OMP provides a superior performance compared to MP, particularly in terms of its convergence property.

The computational complexity of OMP depends on the actual implementation of the least-squares estimation. For example, the QR factorisation approach requires $2Mk + 3M$ operations to obtain the least-squares solution where k is the size of the current support set [30]. It also depends on the number of iterations needed for signal recovery, i.e., in the order of $\mathcal{O}(KMN)$ for K iterations [24, 26, 31]. This complexity is significantly smaller than that of the l_1-minimisation algorithms based on convex optimisation [23, 24, 36]. However, OMP does not offer the strong theoretical guarantees as the l_1-minimisation techniques [23, 31, 36].

Algorithm 2 OMP Algorithm

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$, $\Lambda^{[0]} = \emptyset$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $j^* = \arg\max_j |\mathbf{c}_j|$
    Merge Supports:
      $\Lambda^{[i]} = \Lambda^{[i-1]} \cup j^*$
    Update:
      $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\mathbf{x}\|_2$, where $\mathbf{\Phi}_{\Lambda^{[i]}}$ are the atoms in current set $\Lambda^{[i]}$
      $\mathbf{r}^{[i]} = \mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}}$
  end for
end procedure

It was demonstrated theoretically and empirically in [37] that OMP can reliably recover a K-sparse signal for given $M = O(K \log N)$ noise-free measurements.

Specifically, for a Gaussian sensing matrix (its atoms are drawn independently from the standard Gaussian distribution), OMP can recover any K-sparse signal $\mathbf{x}$ with probability exceeding $(1 - 2\delta)$ given that $\delta \in (0, 0.36)$ and $M \geq CK \log(N/\delta)$ [37]. The constant C satisfies $C \leq 20$ or can be reduced to $C \approx 4$ for a large value of K. However, in contrast to l_1-minimisation, the recovery guarantees of OMP is nonuniform [23, 38, 39]. In particular, when $M = O(K \log N)$, there is high possibility of existence of a K-sparse vector $\mathbf{x}$ for which OMP will select an incorrect element at the first iteration for certain random matrices $\mathbf{\Phi}$ [38, 39]. It has been shown in [38] that uniform recovery guarantees for OMP are impossible for the natural random sensing matrix $\mathbf{\Phi}$.

In addition, the conditions on the sensing matrices required by OMP are more restrictive than the RIP condition [23, 24]. In particular, Tropp [40] showed that the sufficient condition for OMP to recover a K-sparse $\mathbf{x}$ exactly from noise-free measurements is

$$\mu < \frac{1}{2K - 1} \tag{10}$$

where μ is the mutual coherence parameter defined by the maximum value of the modulus of the inner product between two distinct columns of the sensing matrix $\mathbf{\Phi}$ (commonly known as the Mutual Incoherence Property (MIP) introduced in [41]). This result was then extended to the case of noisy measurements in [42]. Note that if MIP holds then RIP also holds, but the converse is not true. Thus the MIP condition is more restrictive than the RIP condition [42]. However, the advantage of the MIP over the RIP is that the MIP is more straightforward to verify for any given matrix $\mathbf{\Phi}$ [42].

Although the RIP has been proved in [43] to hold with high probability for the random construction of the sensing matrix $\mathbf{\Phi}$, there is still probability of failure of the RIP for a particular realisation of a random matrix. In addition, testing whether a given matrix satisfies the RIP is NP-hard and computationally infeasible [44]. In terms of RIP, Wakin and Davenport [45]

Algorithm 3 StOMP and SWOMP Algorithms

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$, $\Lambda^{[0]} = \emptyset$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $J^* = \{j : |\mathbf{c}_j| \geq \lambda^{[i]}\}$
    Merge Supports:
      $\Lambda^{[i]} = \Lambda^{[i-1]} \cup J^*$
    Update:
      $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\mathbf{x}\|_2$
      $\mathbf{r}^{[i]} = \mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}}$
  end for
end procedure

showed that the RIP of order $K + 1$ with

$$\delta_{K+1} < \frac{1}{3\sqrt{K}} \tag{11}$$

is sufficient for OMP to exactly reconstruct any K-sparse signal from noise-free measurements. This condition was then improved to [46]

$$\delta_{K+1} < \frac{1}{\sqrt{K}+1}. \tag{12}$$

2.1.3 Stagewise OMP (StOMP), Stagewise Weak OMP (SWOMP) & Generalised OMP (gOMP)

Although OMP is computationally cheaper than the l_1-minimisation algorithms, it may still not be the best choice for certain class of problems, especially for not very sparse signals and in terms of computational complexity [25, 26, 29, 30]. The main reason for this is the fact that OMP only selects a single column of the sensing matrix $\mathbf{\Phi}$ (a single atom), at each iteration and thus it must run at least as many iterations as the number of nonzero elements in the solution. This computational issue motivates the selection of *multiple* atoms at a time, as has been incorporated in a number of variants of OMP including StOMP [25], SWOMP [29], and gOMP [26]. Although the computational complexity for one iteration of StOMP, SWOMP and gOMP is similar to that of OMP, the StOMP, SWOMP and gOMP algorithms require fewer iterations to obtain the same number of nonzero elements in the solution when compared to OMP. This makes StOMP, SWOMP and gOMP computationally more effective than OMP.

The maximum selection operation

$$\Lambda^{[i]} = \Lambda^{[i-1]} \cup j^*, \quad \text{with} \quad j^* = \arg\max_j |\mathbf{c}_j| \tag{13}$$

in OMP is replaced with a thresholding operation in StOMP and SWOMP as follows:

$$\Lambda^{[i]} = \Lambda^{[i-1]} \cup \{j : |\mathbf{c}_j| \geq \lambda^{[i]}\} \tag{14}$$

Algorithm 4 gOMP Algorithm

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$, L. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$, $\Lambda^{[0]} = \emptyset$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $J^* = \{\text{indices of the } L \text{ largest elements of } \mathbf{c}\}$
    Merge Supports:
      $\Lambda^{[i]} = \Lambda^{[i-1]} \cup J^*$
    Update:
      $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\mathbf{x}\|_2$
      $\mathbf{r}^{[i]} = \mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}}$
  end for
end procedure

where $\lambda^{[i]}$ is the threshold at iteration i. The full algorithms of StOMP and SWOMP are summarised in Algorithm 3. The two variants differ only in the choice for this threshold.

The threshold used in StOMP is given by

$$\lambda^{[i]}_{StOMP} = \frac{t^{[i]} \|\mathbf{r}^{[i-1]}\|_2}{\sqrt{M}} \tag{15}$$

where the typical value of the parameter $t^{[i]}$ is $2 \leq t^{[i]} \leq 3$. Two threshold strategies of $t^{[i]}$ were proposed explicitly in [25] for the case of random sensing matrix $\mathbf{\Phi}$ generated from a uniform spherical ensemble, i.e. the columns of $\mathbf{\Phi}$ are independently and identically distributed (i.i.d.) points on the unit sphere. These strategies were motivated by classical detection criteria of false alarm control and false discovery control. However, theoretical performance guarantees of StOMP are not available for more general sensing matrices $\mathbf{\Phi}$.

Moreover, from the practical perspective, the choice of the parameter $t^{[i]}$ appears critical to the performance of StOMP, but the question on the optimal selection of $t^{[i]}$ has not been fully addressed in the literature [29]. It was demonstrated in [29] that StOMP sometimes terminates prematurely when all inner products, i.e., $|\mathbf{c}_j|$ for all values of j, fall below the threshold $\lambda^{[i]}$. In addition, the simulations in [29] showed mixed results on the performance of StOMP.

The threshold used in SWOMP is given by

$$\lambda^{[i]}_{SWOMP} = \alpha \max_j \{|\mathbf{c}_j|\} \tag{16}$$

where $\alpha \in (0, 1]$ is the 'weakness parameter'. It is important to note that SWOMP enjoys a weakened version of the performance properties of OMP [40].

In contrast, as summarised in Algorithm 4, gOMP selects a fixed number of atoms for each iteration, i.e.

$$\Lambda^{[i]} = \Lambda^{[i-1]} \cup \{J^*(1), \ldots, J^*(L)\} \tag{17}$$

Algorithm 5 ROMP Algorithm

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$, K. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$, $\Lambda^{[0]} = \emptyset$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $J = \{\text{indices of the } K \text{ largest elements of } \mathbf{c}\}$
    Regularise:
      $J^* = \arg\max_{J_l} \|\mathbf{c}_{J_l}\|_2$
      where J_l is all subsets of J with comparable magnitudes:
        $|\mathbf{c}_a| \leq 2|\mathbf{c}_b| \quad \text{for all } a, b \in J_l$
    Merge Supports:
      $\Lambda^{[i]} = \Lambda^{[i-1]} \cup J^*$
    Update:
      $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\mathbf{x}\|_2$
      $\mathbf{r}^{[i]} = \mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}} \hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}}$
  end for
end procedure

where L is the number of atoms selected at each iteration, and $J^*(l)$ is the index of the l-th largest absolute value of the entries of $\mathbf{c}$ explicitly defined as

$$J^*(l) = \underset{j \setminus \{J^*(l-1),\ldots,J^*(1)\}}{\arg\max} |\mathbf{c}_j|. \tag{18}$$

For the noise-free measurement scenario, the RIP order of LK with

$$\delta_{LK} < \frac{\sqrt{L}}{\sqrt{K} + 3\sqrt{L}}, \quad (K > 1) \tag{19}$$

is a sufficient condition for gOMP to obtain an exact recovery of any K-sparse vector within K iterations [26]. A performance bound on the estimation error for the signal reconstruction in the presence of noise was also derived in [26].

2.1.4 Regularised OMP (ROMP)

An alternative modification on the identification step was proposed in the ROMP algorithm [23, 31] which, at each iteration, selects K largest entries of the correlation $\mathbf{c}$ and groups them into subsets J_l with comparable magnitudes:

$$|\mathbf{c}_a| \leq 2|\mathbf{c}_b| \quad \text{for all } a, b \in J_l. \tag{20}$$

ROMP then selects the set J_l with the maximum energy $\|\mathbf{c}_{J_l}\|_2$. The algorithm is summarised in Algorithm 5.

In contrast to OMP and its aforementioned variants, ROMP enjoys uniform recovery guarantees [23]. In particular, the RIP of order $2K$ with

$$\delta_{2K} < 0.03/\sqrt{\log K} \tag{21}$$

is the sufficient condition for ROMP to exactly reconstruct any K-sparse signal $\mathbf{x}$ from its noise-free measurement $\mathbf{y} = \mathbf{\Phi x}$ [23]. For the scenario with the presence of noise $\mathbf{n}$ in the measurement $\mathbf{y} = \mathbf{\Phi x} + \mathbf{n}$, the theoretical performance bound of ROMP is given by

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq 104\sqrt{\log K}\|\mathbf{n}\|_2, \tag{22}$$

as discussed in [31]. In fact, the logarithmic factor in this expression yields a stronger theoretical requirement for the RIP compared to that of the l_1-minimisation methods [29, 31]. However, from a practical point of view, one may be more interested in the average performance than the worse case performance provided by these theoretical analyses. It was empirically demonstrated in [29] that the average performance of ROMP was notably worse than OMP and StOMP.

The computation for one iteration of ROMP is slightly more demanding than that of OMP due to the differences in the identification and regularisation steps. The selection of K largest elements of $\mathbf{c}$ can be done via a sorting algorithm which typically requires a complexity of $\mathcal{O}(N \log N)$ [31]. Since the selected support set J is already sorted, the regularisation step can be performed effectively by searching over consecutive intervals of J, and thus it only requires the complexity of $\mathcal{O}(K)$ [31]. However, ROMP may select multiple atoms at a time. As a result, ROMP may require less run time than OMP if a smaller number of iterations is required for ROMP to achieve the same sparsity level (i.e., same number of non-zero elements) in the solution.

2.1.5 Compressive Sampling MP (CoSaMP) & Subspace Pursuit (SP)

In contrast to the other aforementioned variants of OMP which only focus on the modification to the identification step of OMP, the CoSaMP algorithm proposed in [32] and the SP algorithm proposed in [24] include an additional *pruning* step in each iteration. The CoSaMP and SP algorithms are summarised in Algorithm 6. The main idea behind CoSaMP and SP is that they *maintain a fixed number of nonzero elements in each active set* $\Lambda^{[i]}$ for each iteration by removing insignificant elements via the pruning step.

Note that, in the other greedy algorithms previously described, once an atom is selected, it will always stay in the active set until the algorithm terminates. Specifically, the pruning step in CoSaMP and SP only retains K largest entries in the least-squares optimisation for the merged support set $\Lambda^{[i]}$. CoSaMP and SP then re-estimate the least-squares solution corresponding to the retained support set $\tilde{\Lambda}^{[i]}$. The only difference between CoSaMP and SP is that CoSaMP adds $2K$ new atoms to the active support set in the identification step while SP only adds K new atoms.

The main advantage of these two algorithms is that they are not computationally complex and provide strong theoretical guarantees that are comparable to those derived for the convex l_1-optimisation methods [21, 24, 32]. However, the drawback of CoSaMP and SP is that they require the knowledge of the sparsity level K as an input, which may not be known a priori in practice.

Algorithm 6 CoSaMP and SP Algorithms

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$, K. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$, $\Lambda^{[0]} = \emptyset$, $\tilde{\Lambda}^{[0]} = \emptyset$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $J^* = \{$indices of the $2K$ largest elements of $\mathbf{c}\}$ (CoSaMP)
      or $J^* = \{$indices of the K largest elements of $\mathbf{c}\}$ (SP)
    Merge Supports:
      $\Lambda^{[i]} = \tilde{\Lambda}^{[i-1]} \cup J^*$
    Estimate Least Squares (LS) Solution:
      $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}}\mathbf{x}\|_2$
    Pruning:
      $\tilde{\Lambda}^{[i]} = \{$indices of the K largest elements of $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}}\}$
    Update:
      $\hat{\mathbf{x}}^{[i]}_{\tilde{\Lambda}^{[i]}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{\tilde{\Lambda}^{[i]}}\mathbf{x}\|_2$
      $\mathbf{r}^{[i]} = \mathbf{y} - \mathbf{\Phi}_{\tilde{\Lambda}^{[i]}}\hat{\mathbf{x}}^{[i]}_{\tilde{\Lambda}^{[i]}}$
  end for
end procedure

2.1.6 Gradient Pursuit (GP), Conjugate GP (CGP) & Stagewise Weak CGP(SWCGP)

Another trend of greedy pursuit technique based on directional updates was proposed in [29, 30], namely GP, CGP and SWCGP. The main idea behind these directional algorithms is that they exploit directional optimisation to update the coefficients of the selected elements in each iteration instead of using orthogonal projection, i.e., least-squares estimation, as in OMP and its variations. The replacement of the costly orthogonal projection by the directional optimisation leads to computationally more efficient algorithms while still retaining similar performance as OMP. Directional optimisation refers to an iterative technique of finding a local minimum of a given cost function by starting with an initial point in the parameter space and moving towards the direction that minimises the cost function. A great example of directional optimisation is the gradient method where the update direction is determined by the gradient of the cost function at the current point.

The expression of directional update is [29, 30]

$$\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \hat{\mathbf{x}}^{[i-1]}_{\Lambda^{[i]}} + a^{[i]}\mathbf{d}^{[i]}_{\Lambda^{[i]}} \tag{23}$$

where $\mathbf{d}^{[i]}_{\Lambda^{[i]}}$ is the update direction and $a^{[i]}$ is the step size. The optimal value of the step size $a^{[i]}$ to minimise $\|\mathbf{y} - \mathbf{\Phi}\hat{\mathbf{x}}^{[i]}\|_2^2$ over all $\hat{\mathbf{x}}^{[i]}$ with support $\Lambda^{[i]}$ (i.e., the same quadratic cost as in OMP) is explicitly given by [30, 47]

$$a^{[i]} = \frac{\langle \mathbf{r}^{[i]}, \mathbf{b}^{[i]} \rangle}{\|\mathbf{b}^{[i]}\|_2^2} \tag{24}$$

Algorithm 7 Gradient Pursuit Algorithm

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\mathbf{r}^{[0]} = \mathbf{y}$, $\hat{\mathbf{x}}^{[0]} = \mathbf{0}$, $\Lambda^{[0]} = \emptyset$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Identify:
      $\mathbf{c} = \mathbf{\Phi}^T \mathbf{r}^{[i-1]}$
      $j^* = \arg\max_j |\mathbf{c}_j|$
    Merge Supports:
      $\Lambda^{[i]} = \Lambda^{[i-1]} \cup j^*$
    Directional Update:
      compute $\mathbf{d}^{[i]}_{\Lambda^{[i]}}$ (see text for details)
      $\mathbf{b}^{[i]} = \mathbf{\Phi}_{\Lambda^{[i]}} \mathbf{d}^{[i]}_{\Lambda^{[i]}}$, $a^{[i]} = \frac{\langle \mathbf{r}^{[i]}, \mathbf{b}^{[i]} \rangle}{\|\mathbf{b}^{[i]}\|_2^2}$
      $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}} = \hat{\mathbf{x}}^{[i-1]}_{\Lambda^{[i]}} + a^{[i]} \mathbf{d}^{[i]}_{\Lambda^{[i]}}$
      $\mathbf{r}^{[i]} = \mathbf{r}^{[i-1]} - a^{[i]} \mathbf{b}^{[i]}$
  end for
end procedure

where $\mathbf{b}^{[i]} = \mathbf{\Phi}_{\Lambda^{[i]}} \mathbf{d}^{[i]}_{\Lambda^{[i]}}$ and $\langle \cdot \rangle$ denotes the inner product operation. The pseudocode for directional pursuit family is given in Algorithm 7. Note that MP and OMP can be naturally cast into the framework of directional pursuit with update directions of $\delta_{j^{[i]}}$ and $\mathbf{\Phi}^{\dagger}_{\Lambda^{[i]}} \mathbf{c}_{\Lambda^{[i]}}$ (the superscript $\dagger$ denotes the matrix pseudoinverse operation).

GP utilises the negative gradient of the cost quadratic cost function $\|\mathbf{y} - \mathbf{\Phi}\hat{\mathbf{x}}^{[i]}\|_2^2$ as the update direction [30]

$$\mathbf{d}^{[i]}_{\Lambda^{[i]}} = \mathbf{\Phi}^T_{\Lambda^{[i]}} (\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}} \hat{\mathbf{x}}^{[i-1]}_{\Lambda^{[i]}}) = \mathbf{c}_{\Lambda^{[i]}} \tag{25}$$

where $\mathbf{c}_{\Lambda^{[i]}}$ is a subvector of $\mathbf{c}$ with support $\Lambda^{[i]}$.

On the other hand, the directional update in CGP exploits the conjugate gradient method which is widely used to solve quadratic optimisation problems [47]. Fundamentally, to minimise a cost function of $(1/2)\mathbf{x}^T\mathbf{G}\mathbf{x} - \mathbf{y}^T\mathbf{x}$ (i.e., equivalent to solving $\mathbf{y} = \mathbf{G}\mathbf{x}$ for $\mathbf{x}$), the conjugate gradient method successively applies line minimisations along directions which are G-conjugate, where set of directions $\{\mathbf{d}^{[1]}, \mathbf{d}^{[2]}, \ldots, \mathbf{d}^{[i]}\}$ is defined as G-conjugate if $\langle \mathbf{d}^{[i]}, \mathbf{G}\mathbf{d}^{[j]} \rangle = 0$ for all $i \neq j$. Note that $\langle \cdot, \cdot \rangle$ denotes the inner product between two vectors. In the framework of directional pursuit, CGP aims to minimise the cost function of $\|\mathbf{y} - \mathbf{\Phi}_{\Lambda^{[i]}} \mathbf{x}_{\Lambda^{[i]}}\|_2^2$ by calculating an update direction that is $\mathbf{G}_{\Lambda^{[i]}}$-conjugate to all previous update directions. Here, $\mathbf{G}_{\Lambda^{[i]}} = \mathbf{\Phi}^T_{\Lambda^{[i]}} \mathbf{\Phi}_{\Lambda^{[i]}}$.

The main advantage of the conjugate gradient method is that the conjugate directions can be computed iteratively using

$$\mathbf{d}^{[i]} = \mathbf{c}^{[i]} + \beta^{[i]} \mathbf{d}^{[i-1]} \tag{26}$$

if the first conjugate direction is initialised to $\mathbf{d}^{[1]} = -\mathbf{c}^{[1]}$. Here, $\beta^{[i]}$ is given by $\beta^{[i]} = \langle \mathbf{c}^{[i]}, \mathbf{G}\mathbf{d}^{[i-1]} \rangle / \langle \mathbf{d}^{[i-1]}, \mathbf{G}\mathbf{d}^{[i-1]} \rangle$ to ensure $\langle \mathbf{d}^{[i]}, \mathbf{G}\mathbf{d}^{[i-1]} \rangle = 0$. It is important to note that this procedure leads to $\mathbf{d}^{[i]}$ conjugate to all previous directions $\mathbf{d}^{[1]}, \ldots, \mathbf{d}^{[i-1]}$. This principle is adopted to compute the update direction of CGP, i.e.,

$$\mathbf{d}^{[i]}_{\Lambda^{[i]}} = \mathbf{c}^{[i]}_{\Lambda^{[i]}} + \beta^{[i]} \mathbf{d}^{[i-1]}_{\Lambda^{[i]}} \tag{27}$$

where

$$\beta^{[i]} = \frac{\left\langle (\mathbf{\Phi}_{\Lambda^{[i-1]}} \mathbf{d}^{[i-1]}_{\Lambda^{[i-1]}}), (\mathbf{\Phi}_{\Lambda^{[i]}} \mathbf{c}_{\Lambda^{[i]}}) \right\rangle}{\| \mathbf{\Phi}_{\Lambda^{[i-1]}} \mathbf{d}^{[i-1]}_{\Lambda^{[i-1]}} \|_2^2}. \tag{28}$$

However, in contrast to that in (26), the iterative procedure in (27) only guarantees the conjugacy of $\mathbf{d}^{[i]}_{\Lambda^{[i]}}$ to the last preceding direction $\mathbf{d}^{[i-1]}_{\Lambda^{[i-1]}}$ but not to all the previous directions because the dimensionality of the solution space varies over iterations. Therefore, using (27) leads to an approximate version of CGP (referred to as CGP throughout this report for simplicity). The more sophisticated version of CGP can be found in [29], where each new direction is calculated to ensure the conjugacy with all previous directions. However, this version is computationally more complex than the approximated CGP.

In common with MP, the identification step of GP requires a matrix vector multiplication and N operations. Computing the size step for the directional update of GP requires an extra matrix vector multiplication and $2M$ operations [30]. In addition, updating $\hat{\mathbf{x}}^{[i]}_{\Lambda^{[i]}}$ takes k operations where k is the size of the current support set, and updating the residual $\mathbf{r}^{[i]}$ can be done via M operations. On the other hand, compared to GP, CGP requires extra computations to compute the update direction $\mathbf{d}^{[i]}$ including one additional matrix vector multiplication and $M + k$ additional operations [30].

Similar to MP and OMP, GP and CGP only selects a single element at each iteration making it not suitable for large-scale problems in terms of computational performance. Motivated by this issue, a stagewise weak version of CGP (SWCGP) was proposed in [30] allowing multiple elements to be selected at each iteration. Since the extension of CGP to SWCGP is analogous with the extension of OMP to SWOMP which has been discussed in Section 2.1.3, the details of SWCGP is omitted here. The interested readers may refer to [30] for more detailed description of SWCGP.

2.1.7 Stopping Criteria for Greedy Algorithms

Several criteria can be used as halting rules for the greedy algorithms. The first option is to stop the greedy algorithms when the l_2-norm of the residual $\mathbf{r}^{[i]}$ falls below a preset threshold. This criterion aims to achieve a certain bound on the reconstruction error.

The second criterion is based on the sparsity level of the solution vector $\hat{\mathbf{x}}^{[i]}$, i.e., the number of nonzero elements in the solution vector. This stopping criterion is normally used when the sparsity level of the actual solution is a priori known or a fixed number of elements is desired to approximate the solution.

The greedy algorithms can also be halted when the change in the l_2-norm of residual or the maximum correlation between the residual and the atoms fall below some threshold. These two criteria ensure the greedy algorithms achieve a certain bound on the reconstruction error. Depending on the applications, one or more criteria can be applied to the greedy algorithms.

Algorithm 8 IRLS Algorithm

procedure INPUT: $\mathbf{y}$, $\mathbf{\Phi}$, K. OUTPUT: $\mathbf{r}^{[i]}$, $\hat{\mathbf{x}}^{[i]}$
  Initialisations: $\hat{\mathbf{x}}^{[0]} = \mathbf{\Phi}^{\dagger}\mathbf{y}$, $\vartheta^{[1]} = 1$.
  for iteration $i = 1$; $i := i + 1$ until stopping criterion is met **do**
    Compute Weights:
      $w_j^{[i]} = ((|\hat{\mathbf{x}}_j^{[i-1]}|)^2 + \vartheta^{[i-1]})^{p/2-1}$ $(j = 1, \ldots, N)$
    Solution for l_2-norm Objective:
      $\hat{\mathbf{x}}^{[i]} = \mathbf{Q}^{[i]}\mathbf{\Phi}^T(\mathbf{\Phi}\mathbf{Q}^{[i]}\mathbf{\Phi}^T)^{-1}\mathbf{y}$
      where $\mathbf{Q}^{[i]}$ is the diagonal matrix with entries of $1/w_j^{[i]}$
    Updating the Regularisation Parameter:
    if $(\|\hat{\mathbf{x}}^{[i]}\|_2 - \|\hat{\mathbf{x}}^{[i-1]}\|_2) < \sqrt{\vartheta^{[i]}}/100$ **then**
      $\vartheta^{[i+1]} = \vartheta^{[i]}/10$
    else
      $\vartheta^{[i+1]} = \vartheta^{[i]}$
    end if
  end for
end procedure

2.2 l_1-Minimisation Algorithms

The l_1-minimisation approach with the formulation defined in (5) provides a powerful framework for sparse signal recovery with strong theoretical performance guarantees. If the sensing matrix $\mathbf{\Phi}$ satisfies a certain restricted isometry property (RIP), a stable solution of the sparse signal recovery problem is guaranteed to be obtained through the l_1-minimisation. For the case of noise-free signal recovery with $\mathcal{B}(\mathbf{y}) = \{\mathbf{x} : \mathbf{\Phi}\mathbf{x} = \mathbf{y}\}$, the l_1-minimisation can recover any K-sparse signal $\mathbf{x}$ exactly from as few as $O(K\log(N/K))$ measurements if the sensing matrix $\mathbf{\Phi}$ satisfies the RIP of order $2K$ with $\delta_{2K} < \sqrt{2} - 1$ [17, 18]. For the case of noisy signal recovery with $\mathcal{B}(\mathbf{y}) = \{\mathbf{x} : \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2 < \epsilon\}$, the accuracy of the solution $\hat{\mathbf{x}}$ to (5) is bounded by

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq C_0 \frac{\|\mathbf{x} - \mathbf{x}_K^*\|_1}{\sqrt{K}} + C_1\epsilon, \tag{29}$$

where

$$C_0 = 2\frac{1 - (1 - \sqrt{2})\,\delta_{2K}}{1 - (1 + \sqrt{2})\,\delta_{2K}}, \quad C_1 = 4\frac{\sqrt{1 + \delta_{2K}}}{1 - (1 + \sqrt{2})\,\delta_{2K}}, \tag{30}$$

if the sensing matrix $\mathbf{\Phi}$ satisfies the RIP of order $2K$ with $\delta_{2K} < \sqrt{2} - 1$ [18]. Here $\mathbf{x}_K^*$ denotes the best K-sparse approximation of $\mathbf{x}$.

Along with provable performance guarantees, the l_1-minimisation formulation (5) is a convex optimisation problem which generally can be solved effectively via any general-purpose convex optimisation techniques [48, 49]. Specifically, the l_1-minimisation can be posed as a linear program for $\mathcal{B}(\mathbf{y}) = \{\mathbf{x} : \mathbf{\Phi}\mathbf{x} = \mathbf{y}\}$, or it can be considered as a convex program with a conic constraint for $\mathcal{B}(\mathbf{y}) = \{\mathbf{x} : \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2 < \epsilon\}$.

In addition to the l_1-minimisation formulation, there exists the *unconstrained* l_1-minimisation

formulation which has also received great attention in the literature (see [20, 50–53]):

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^N} \left\{ \frac{1}{2}\|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1 \right\}, \tag{31}$$

which is commonly referred to as the l_1-penalty formulation. This l_1-penalty formulation is the core of the well-known basis pursuit denoising (BPDN) algorithm proposed in [53]. In fact, the use of l_1-penalty has a long history outlined in [20]. Here the parameter λ controls the tradeoff between the approximation error and the sparsity of the approximation vector. In fact, the l_1-penalisation formulation (31) is equivalent to the l_1-minimisation formulation (5) with $\mathcal{B}(\mathbf{y}) = \{\mathbf{x} : \|\mathbf{y}-\mathbf{\Phi}\mathbf{x}\|_2 < \epsilon\}$ for some particular value of λ [54]. However, the value of λ, which yields the equivalence between these formulations, is generally unknown. Several methods for determining λ are available in [53, 55, 56]. Moreover, the l_1-penalty optimisation (31) is also equivalent to

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^N} \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \leq t, \tag{32}$$

for appropriate values of λ. Here, t is a positive parameter which is inversely related to λ. In contrast to the l_1-minimisation which is a quadratically constrained linear program, the optimisation in (32) is a quadratic program. This formulation in fact is used in the least absolute shrinkage and selection operator (LASSO) approach [52].

2.3 l_p-Minimisation Algorithms with $p < 1$

The l_p-minimisation approach aims to minimise the l_p-norm ($p < 1$) of the estimate $\hat{\mathbf{x}}$,

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^N} \|\mathbf{x}\|_p \quad \text{subject to} \quad \mathbf{x} \in \mathcal{B}(\mathbf{y}). \tag{33}$$

In fact, the works in [33–35] focused on the case of noise-free measurements $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$, i.e., $\mathcal{B}(\mathbf{y}) = \{\mathbf{x} : \mathbf{\Phi}\mathbf{x} = \mathbf{y}\}$. It is empirically demonstrated in [33] that exact recovery of sparse signals can be achieved with substantially fewer measurements by replacing the l_1-norm by the l_p-norm with $p < 1$. In addition, the theoretical RIP condition of $\mathbf{\Phi}$ for the l_p-minimisation to produce an exact reconstruction of $\mathbf{x}$ is [33]

$$\delta_{aK} + b\delta_{(a+1)K} < b - 1 \tag{34}$$

where

$$b > 1, a = b^{p/(2-p)}, \text{ with } p \in (0, 1]. \tag{35}$$

In fact, this is the generalisation of the RIP condition for $p = 1$ and $b = 3$ presented in [17]. More importantly, this result implies that a weaker condition for exact recovery is obtained for smaller p [33]. Moreover, for the case of random Gaussian sensing matrix $\mathbf{\Phi}$, the l_p-minimisation with $p \in (0, 1]$ can recover exactly any K-sparse signal $\mathbf{x}$ with probability exceeding $1 - 1/\binom{N}{K}$ given that [34]

$$M \geq C_1(p)\, K + p\, C_2(p)\, K \log(N/K), \tag{36}$$

where the constant C_1 and C_2 are determined explicitly and are bounded in p.

The l_p-minimisation with $p < 1$ is a non-convex problem which is intractable as a direct problem, as described in the literature [34]. However, when it is cast as a re-weighted least squares (IRLS) problem, it can be solved via an iterative approach [34, 35]. The main idea behind IRLS is that it replaces the l_p objective function in (33) by a weighted l_2-norm:

$$\hat{\mathbf{x}}^{[i]} = \arg\min_{\mathbf{x}\in\mathbb{C}^N} \sum_{j=1}^{N} w_j^{[i]} \mathbf{x}_j^2 \quad \text{subject to} \quad \mathbf{\Phi x} = \mathbf{y} \tag{37}$$

where the weights $w_j^{[i]}$ $(j = 1, \ldots, N)$ are computed from $\hat{\mathbf{x}}^{[i-1]}$ obtained from the previous iteration. The l_2 objective in (37) becomes the first-order approximation of the l_p objective in (33) when $w_j^{[i]} = |\hat{\mathbf{x}}_j^{[i-1]}|^{p-2}$. However, a regularisation parameter ϑ is introduced to improve the estimation performance [34, 35]:

$$w_j^{[i]} = \left[(|\hat{\mathbf{x}}_j^{[i-1]}|)^2 + \vartheta^{[i-1]} \right]^{p/2-1}. \tag{38}$$

The full description of the algorithm is given in Algorithm 8. Note that this ϑ-regularised IRLS algorithm becomes the FOCUSS algorithm in [57] if the ϑ-regularisation strategy is removed. However, numerical results in [35] show that the performance of the ϑ-regularised IRLS algorithm is significantly superior over the performance of the FOCUSS algorithm.

3 Comparative Performance Study

In this section, we focus our attention on the greedy pursuit family and present a performance comparison of different greedy pursuit algorithms in which the dictionary is built from tilted-wire atoms—the exact solutions of the tilted-wire scatterer model.

3.1 Signal Analysis of Rotating Blades

The scattered signal $s(t)$ in the time domain t received from a rotating blade can be decomposed as the weighted sum of K tilted-wire components [13, 14],

$$s(t) = \sum_{k=1}^{K} \rho_k \, g(t; \boldsymbol{\vartheta}_k), \tag{39}$$

where

$$\begin{aligned} g(t; \boldsymbol{\vartheta}_k) &= g(t; r_k, \psi_k, L_k, \alpha_k) \\ &= A \operatorname{sinc}\left\{ \frac{bL_k}{2} \sin(\Omega t + \psi_k + \alpha_k) \right\} \exp\{ibr_k \sin(\Omega_k t + \psi_k)\} \end{aligned} \tag{40}$$

is the tilted-wire scatterer model of the received signal from a tilted straight wire of length L_k with radial distance r_k to its mid-point from the origin as depicted in Figure 1, while ρ_k is its (generally complex-valued) coefficient. The wire rotates around the origin with angular velocity Ω from initial angle ψ_k, and is tilted at a fixed angle α_k. In (40), A is the normalisation

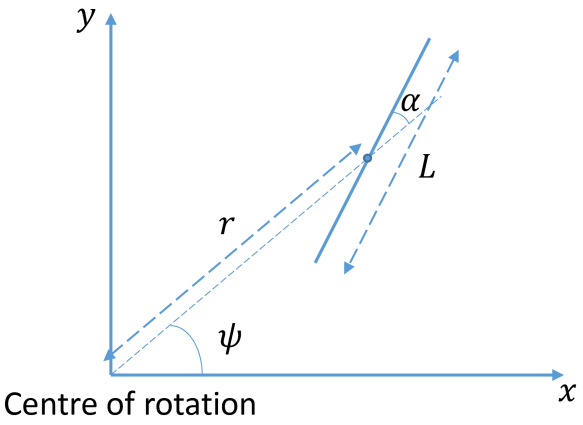


Figure 1: Tilted wire geometry (radar is located in the positive-y direction.)

constant, and the parameter b is defined by $b = 4\pi/\lambda$ with λ being the wavelength of the radar signal.

In the context of compressed sensing, we refer to $g(t;\boldsymbol{\vartheta}_k)$ as a 'tilted-wire atom'. Note that sinc() is the unnormalized sinc function, and $i = \sqrt{-1}$. Writing (39) in vector form yields

$$\mathbf{s} = \sum_{k=1}^{K} \rho_k \, \mathbf{g}_k, \tag{41}$$

where

$$\mathbf{s} = [s(t_1), \ldots, s(t_N)]^T, \text{ and } \mathbf{g}_k = [g(t_1, \boldsymbol{\vartheta}_k), \ldots, g(t_N, \boldsymbol{\vartheta}_k)]^T \tag{42}$$

are vectors of N discrete-time samples of $s(t)$ and $g(t, \boldsymbol{\vartheta}_k)$ respectively. Since an object can be represented as a small number of scattering elements, the problem of radar imaging for rotating blades can be cast into the sparse signal representation problem over an over complete library of tilted-wire atoms. Specifically, the aim is to find a sparse solution of the coefficient vector

$$\boldsymbol{\rho} = [\rho_1, \ldots, \rho_m]^T \tag{43}$$

from

$$\mathbf{s} = \mathbf{G}\boldsymbol{\rho}, \tag{44}$$

where the dictionary

$$\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_M] \tag{45}$$

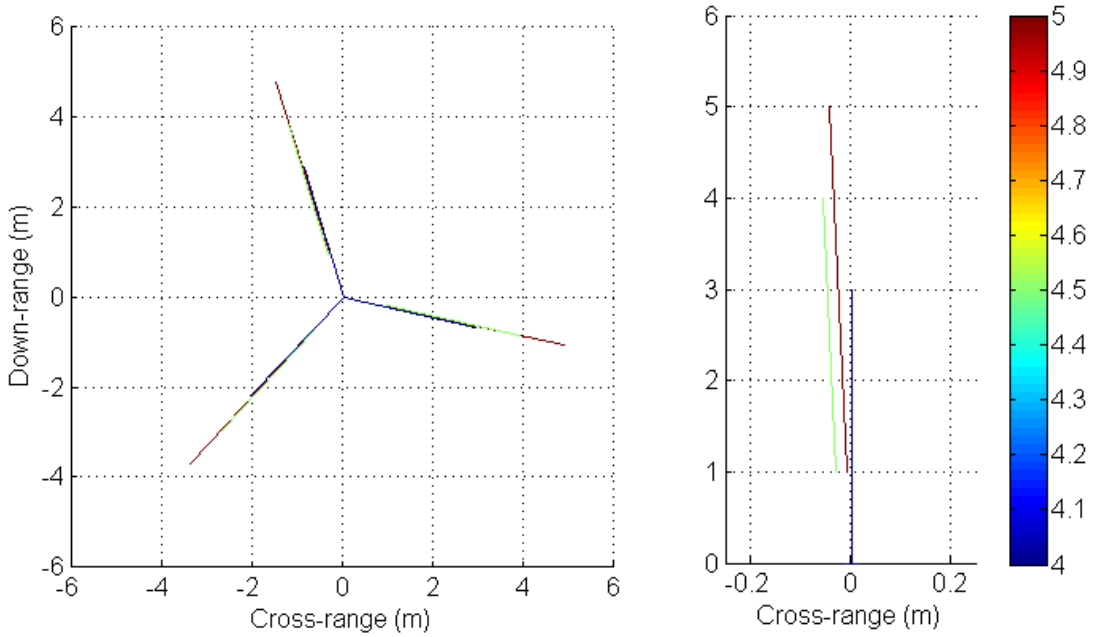
consists of M tilted-wire atoms spanning over a discrete (generally 4-dimensional) grid of parameters $\boldsymbol{\vartheta}_k$. Note that, since the scattering characteristics from the approaching edges (i.e., resulting positive Doppler frequency) are different to those from the receding edges, the positive and negative Doppler flashes are processed separately [14]. This can be done by including both positive and negative radius values [14].

3.2 Results With Simulated Data

Consider a simulated helicopter rotor system with three blades rotating at 40 rad/s, with each blade consisting of three wires with parameters given in Table 1. The parameter grids used for constructing the function dictionary $\mathbf{G}$ are $1 \leq L \leq 5$ m in steps of 0.5 m, $-4 \leq r \leq 4$ m in steps of 0.5 m, $-1^\circ \leq \alpha \leq 1^\circ$ in steps of 0.5°, and ψ spanning over a interval of $\pm 5^\circ$ in steps of 0.5° around -12.5°, -132.5° and 107.5°, which are the angular positions ψ of the blades when they are orthogonal to the radar. The wire parameters have been selected such that they are relatively very close in the physical parameter space, resulting in a significant level of correlation among 'nearby' atoms in the dictionary.

Table 1: Tilted-wire parameters used for the target model.

Blade	Wire	Weighting Coeff.	ψ (deg)	α (deg)	r (m)	L (m)
	1	4.5	-12	-0.5	2.5	3
1	2	5	-12.5	0	3	4
	3	4	-13	0	1.5	3
	1	4.5	-132	-0.5	2.5	3
2	2	5	-132.5	0	3	4
	3	4	-133	0	1.5	3
	1	4.5	108	-0.5	2.5	3
3	2	5	107.5	0	3	4
	3	4	107	0	1.5	3



(a) Orientation of the three wires

(b) Exploded view of a single wire

Figure 2: Scatter plots of the target model showing the orientation and length of the 3 tilted wires. Color bar indicates the amplitude of the wire coefficients.

Figure 2 is a plot of the 3-bladed rotor showing the orientation and length of each wire. The synthetic data for the received signal, obtained by a pulse-Doppler radar operating at the transmitted frequency of 9.5 GHz and the pulse repetition frequency (PRF) of 66 kHz, is generated using (39) for one cycle of rotor rotation. The signal-to-noise ratio (SNR) is set to 20 dB. The time-domain and spectrogram plots of the synthetic received signal are given in Figure 3.

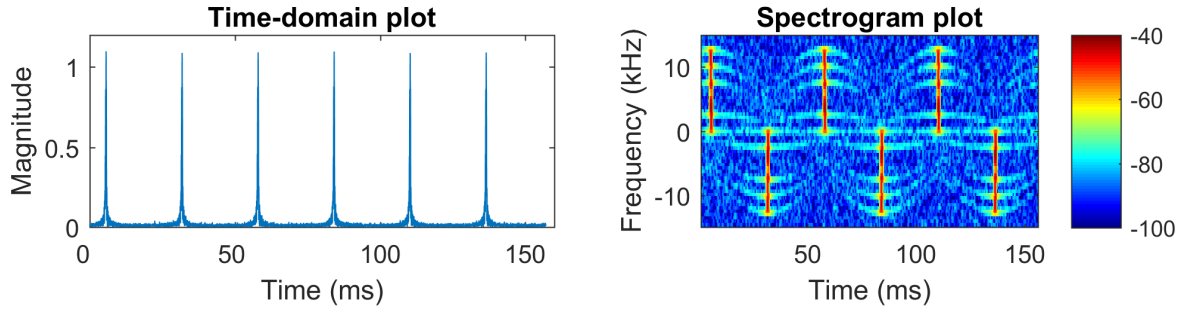


Figure 3: Time-domain and spectrogram plots of the synthetic signal.

Figure 4 shows the comparison in the time domain of the reconstructed signals obtained by MP, OMP, CGP, gOMP, CoSaMP and ROMP for a single Monte Carlo (MC) run. In the frequency domain, the techniques appear almost the same, with no physically significant differences. Figure 5 provides the corresponding scatter plots of all tilted-wire atoms found in the signal representation. Taking into account the relative magnitudes of representing atoms, again the techniques produce very similar results. Here, the results in Figures 4-5 are obtained using 12 atoms per blade flash. These results indicate, at least qualitatively, that even when the scatterers are very close together in the spatial parameter space, the MP, OMP, CGP, and CoSaMP techniques still perform quite satisfactorily at least in term of representing the original signal in time domain.

Figures 6–8 show the root-mean-squared-error (RMSE) of the reconstruction signal and the averaged running time versus the sparsity level in the solution vector for MP, OMP, CGP, gOMP, CoSaMP and ROMP. The RMSE of the reconstructed signal is computed over short intervals with 51 samples around the six main blade flashes and is normalised by the l_2-norm of the original signal in the corresponding intervals. The sparsity level in the solution vector means the number of collected atoms for each blade flash. Here, the results in Figures 6–8 were obtained using 50 MC runs.

A general indication observable from these results is as follows: OMP is the most accurate (smallest reconstruction error in a least-squares sense) but most computationally expensive. CGP is very similar to OMP in terms of accuracy performance, especially for high sparsity levels, and with approximately one third the computational cost of OMP. MP is simple, as fast as CGP, but with an error of about 70 or 80 percent larger than that of OMP. The other techniques are as fast as CGP but with excessive errors, at least for this particular simulated target model.

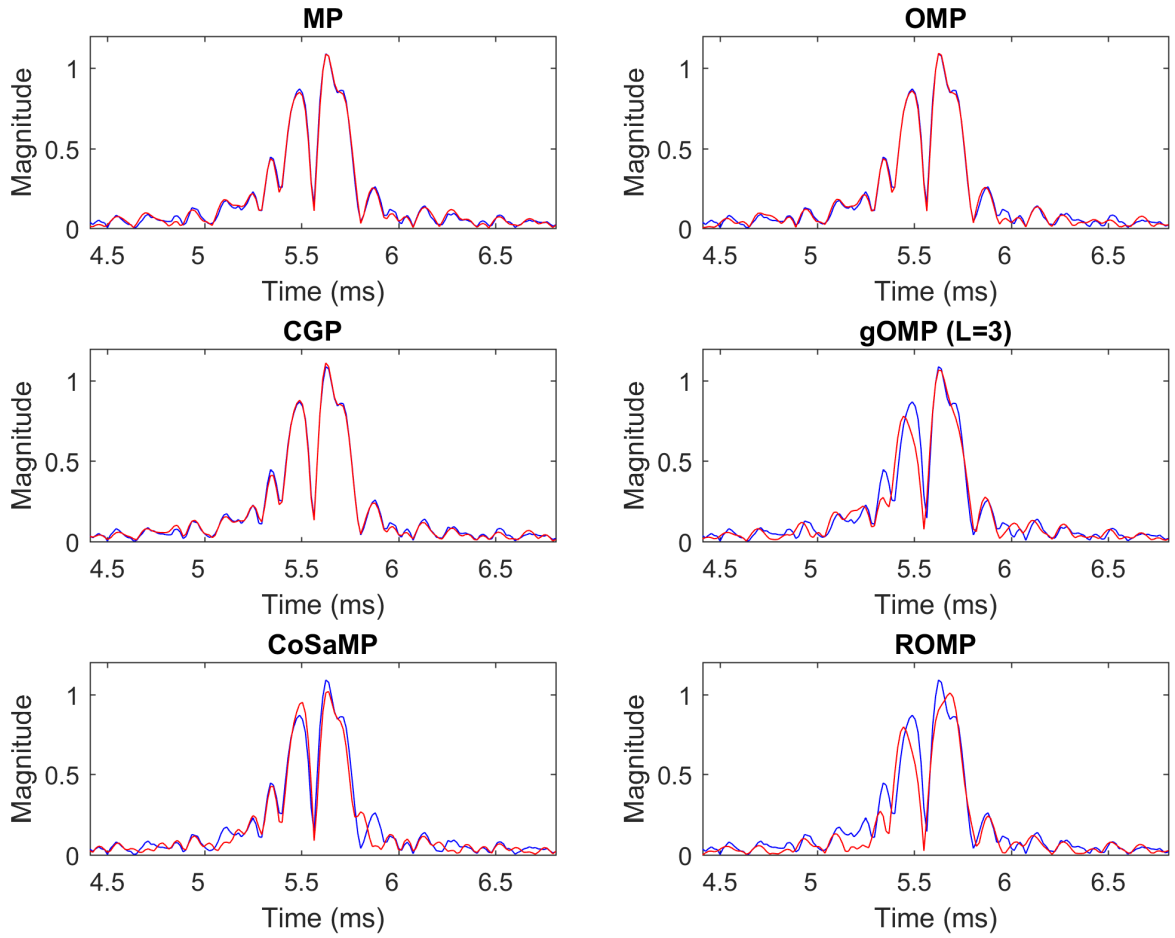


Figure 4: Comparison in time domain of the reconstructed signals (red trace) and the original signal (blue trace) around the first blade flash of the simulated data, for each of the greedy algorithms.

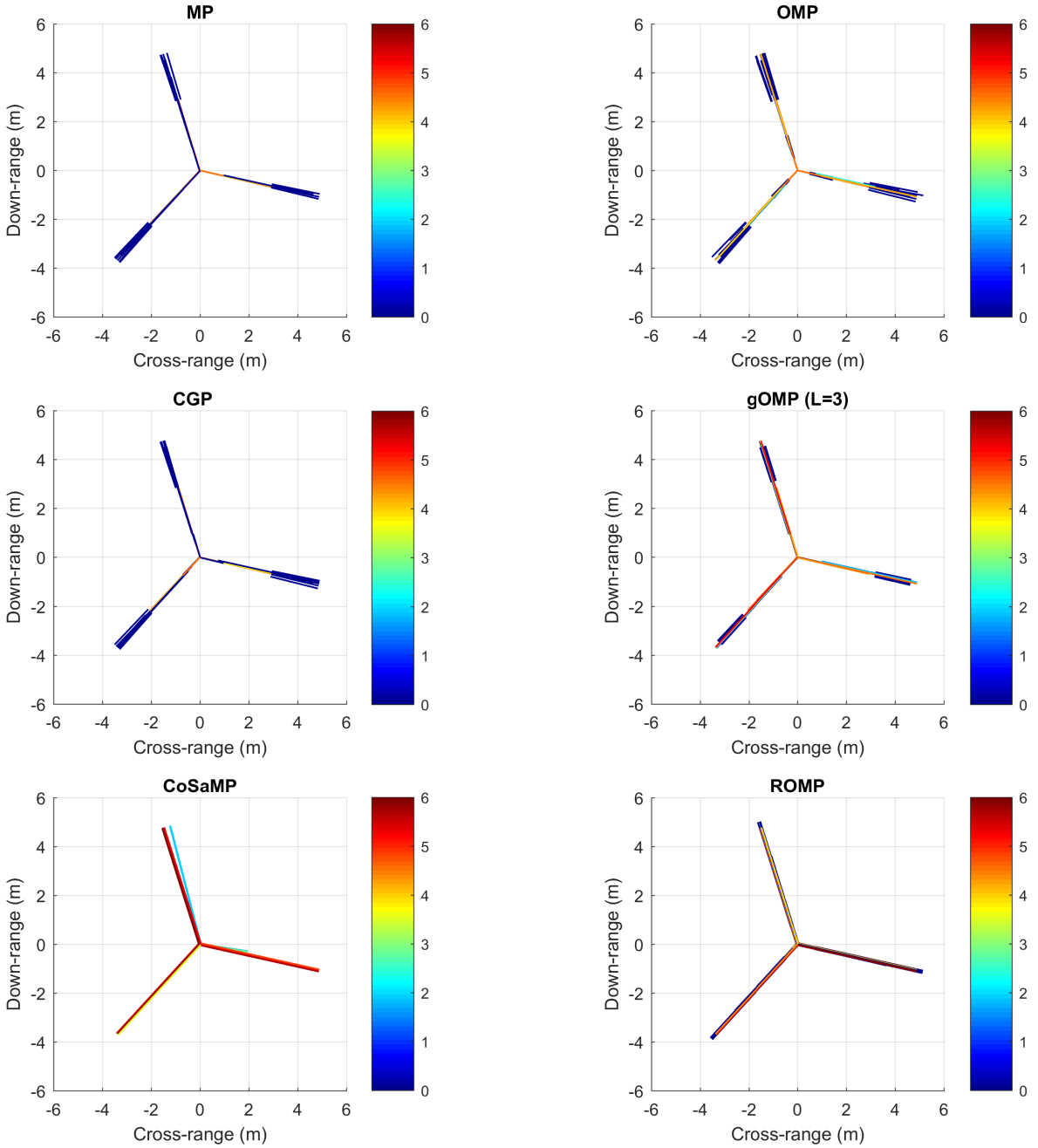


Figure 5: Scatter plots of the extracted tilted-wire atoms representing the simulated signal, with each of the greedy algorithms. Color bar indicates the atom amplitude.

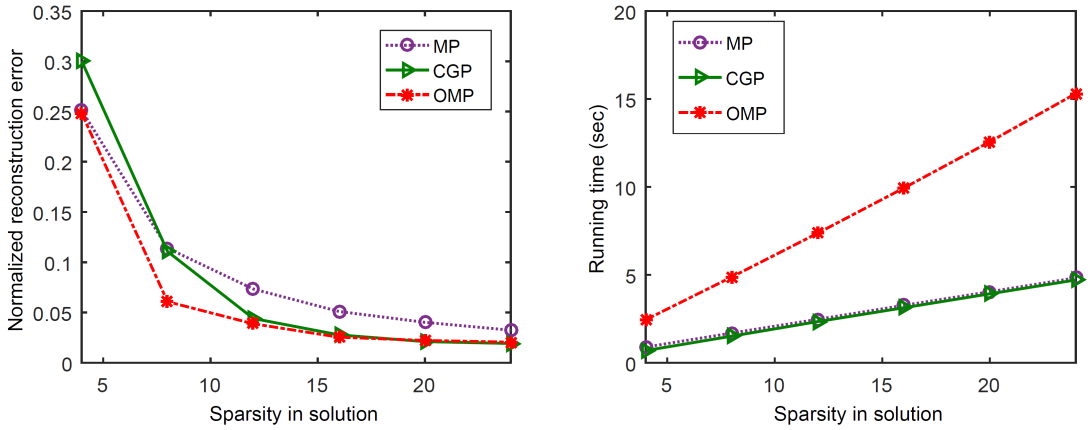


Figure 6: Normalised reconstruction error and running time after processing synthetic data using MP, OMP and CGP.

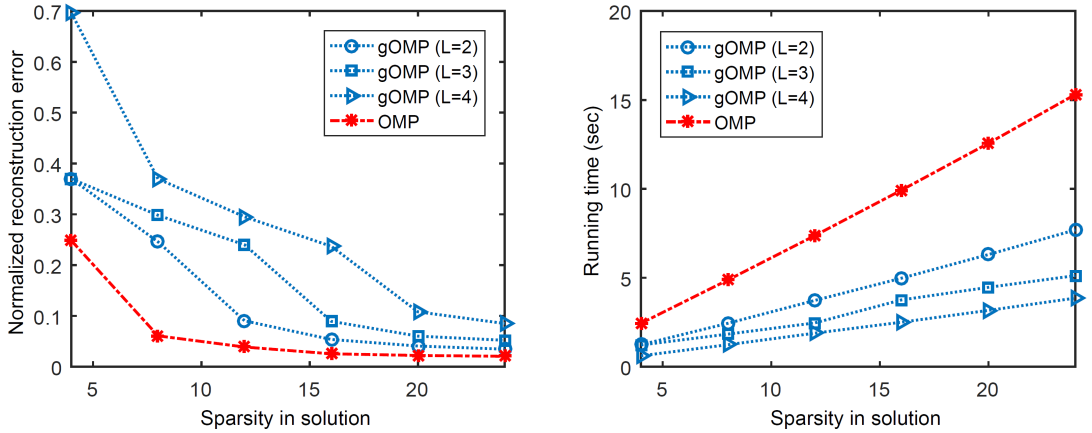


Figure 7: Normalised reconstruction error and running time after processing synthetic data using gOMP.

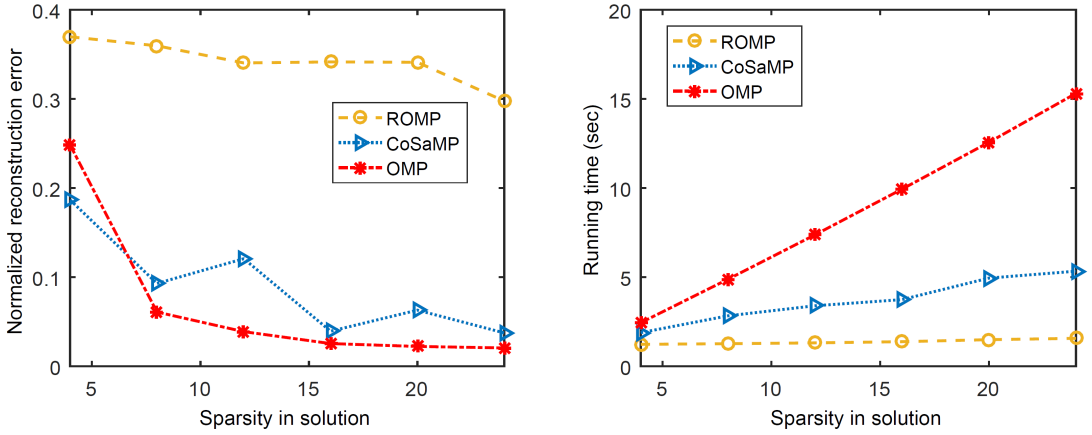


Figure 8: Normalised reconstruction error and running time after processing synthetic data using ROMP, CoSaMP and OMP.

3.3 Results With Real Data

This section uses real data from a 3-bladed Squirrel helicopter collected with an experimental radar operating at the frequency of 9.5 GHz and the PRF of 66 kHz. The raw signal is high-pass filtered to attenuate low-frequency components. The rotation speed and the number of rotor blades are estimated based on [58]. A rough estimate of the initial angles of the blades are also obtained based on the location of the main blade flashes in the time-domain signal. Figure 9 displays the time-domain and spectrogram plots of the preprocessed signal. Though the pre-processed signal contains all return components from the helicopter which include the main rotor, tail rotor, rotor hub, and the fuselage (or aircraft body), only the main rotor blades will be analysed in this study.

The function dictionary used here is constructed from a parameter grid for ψ spanning over interval of $\pm 5^\circ$ in steps of 0.5° around these initial angle estimates. The same parameter grids for L, r and α are used as in Section 3.2 for the simulated data example. Here, a coherent processing interval of 150 ms is used corresponding to approximately one cycle of rotor revolution.

Figures 10 and 11 show the time-domain and frequency-domain plots of the reconstructed signals obtained by MP, OMP, CGP, gOMP, CoSaMP and ROMP as compared to the original real signal. The agreement is reasonable though not as close as in the simulated example, which can be attributed to various real effects: a real rotor blade in flight deviates signifantly from the straight and static shape – it flaps and bends in a rather random and chaotic manner, as can be verified elsewhere [59]. Another interesting observation is that the components in lower Doppler frequency regions near the rotor hub were picked up by MP, OMP, and CGP, but not by the other techniques. Note that MP, OMP, and CGP pick up only one atom in the Identify step of each iteration, whereas the other techniques pick up multiple atoms.

Figure 12 provides the corresponding scatter plots showing the wire model parameters for all the collected atoms. Again, the results in Figures 10-12 are obtained using 12 atoms per flash. The relative accuracies of these techniques can be qualitatively assessed through the locations of the wire elements: MP, OMP, and CGP produce reasonable wire-frame images, while CoSaMP is the worst.

The quantitative results in Figures 13–15 confirm the above qualitative assessment. As before, these plots show the normalised l_2-norm of the reconstruction error and the running time versus the sparsity level in the solution vector for MP, OMP, CGP, gOMP, CoSaMP and ROMP. The l_2-norm of the reconstruction error is computed over short intervals with 51 samples around the six main blade flashes and is normalised by the l_2-norm of the original signal in the corresponding intervals. The sparsity level in the solution vector means the number of collected atoms for each flash.

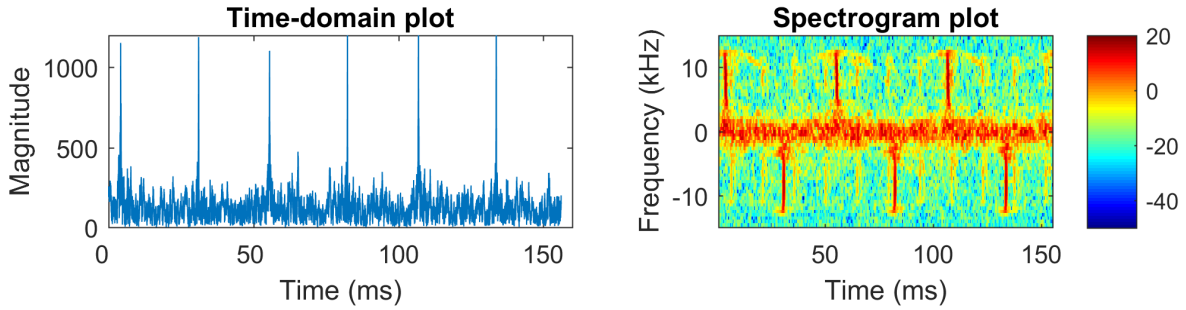


Figure 9: Time-domain and spectrogram (frequency-domain) plots of the Squirrel helicopter data.

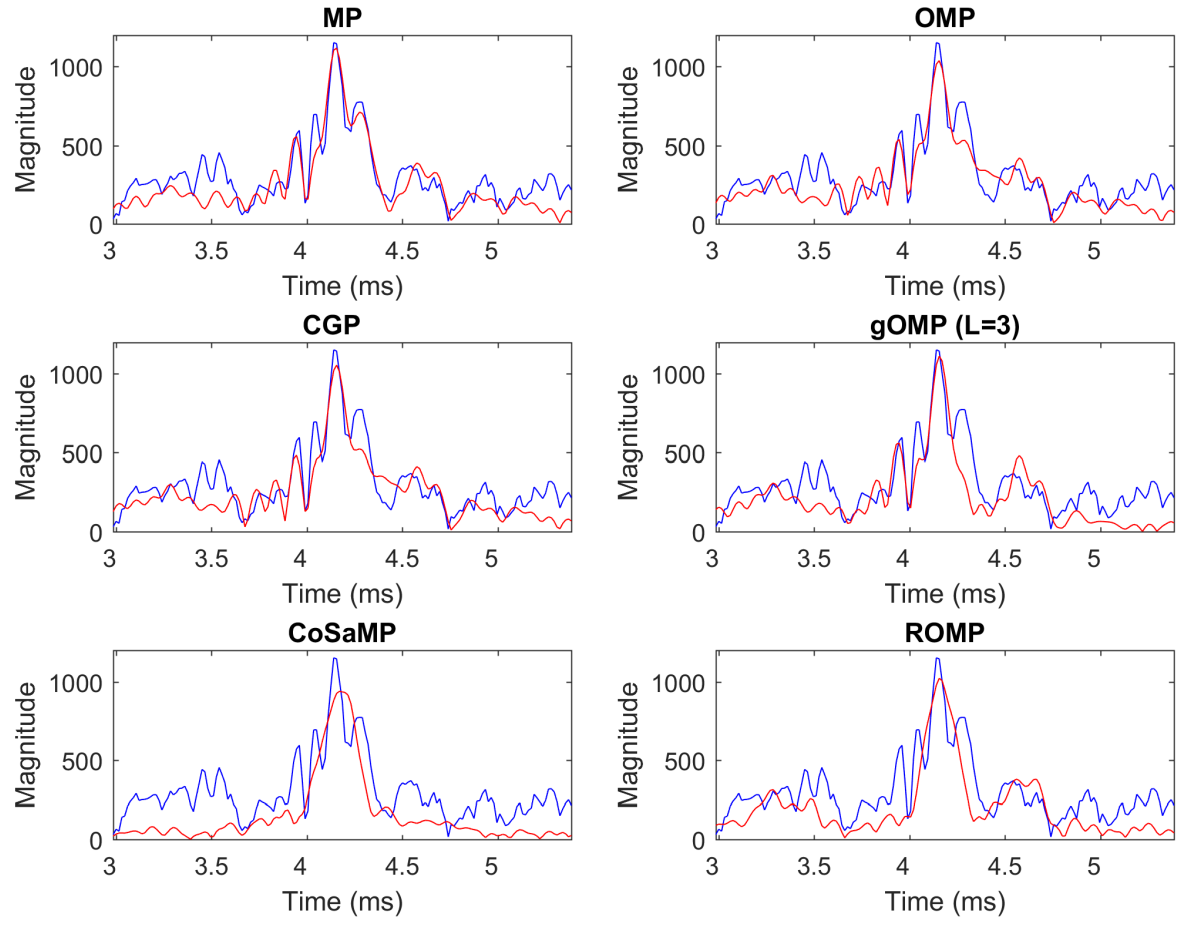


Figure 10: Comparison in the time domain of the reconstructed signals (red line) and the original signal (blue line) around the first blade flash in real data, for each of the greedy algorithms.

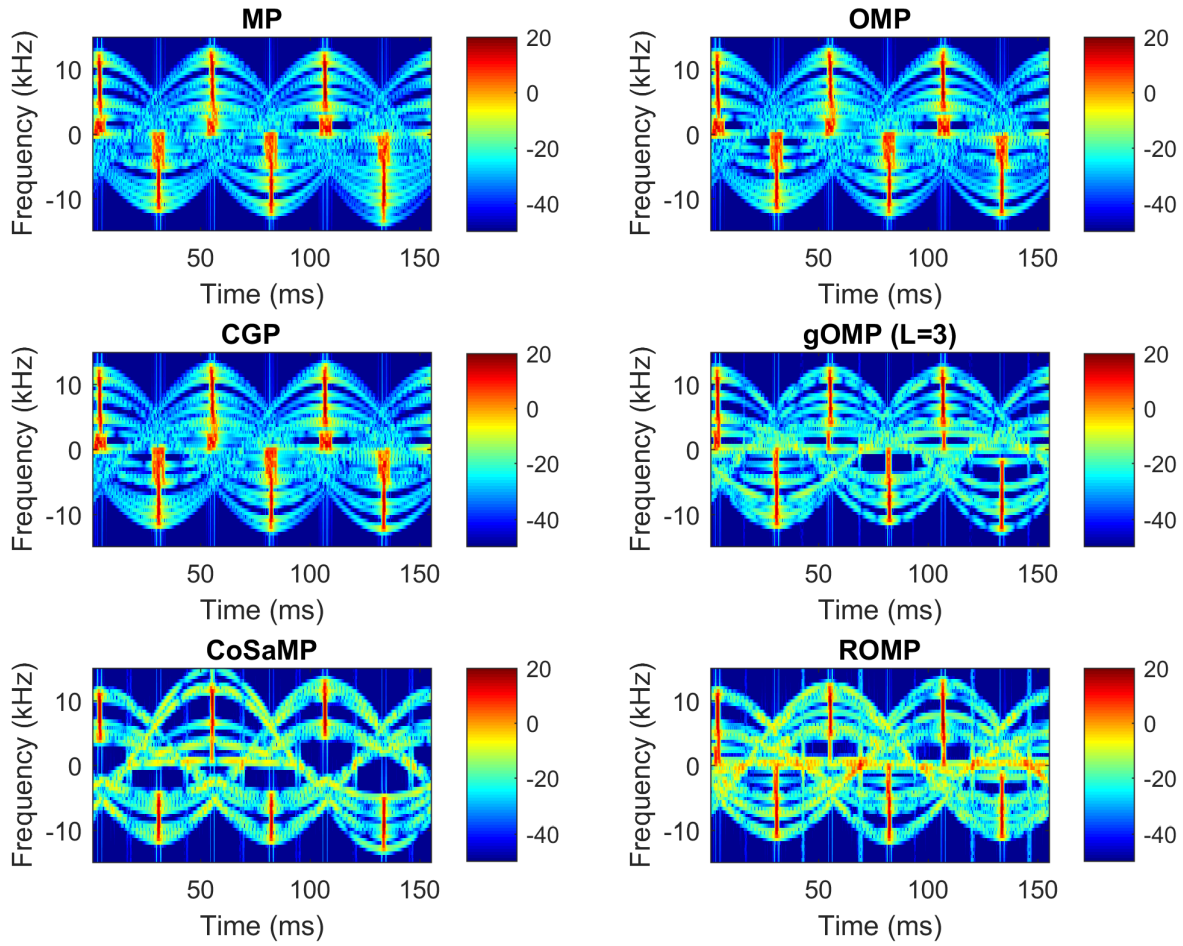


Figure 11: Spectrogram plots of the reconstructed signals after processing real data with each of the greedy algorithms.

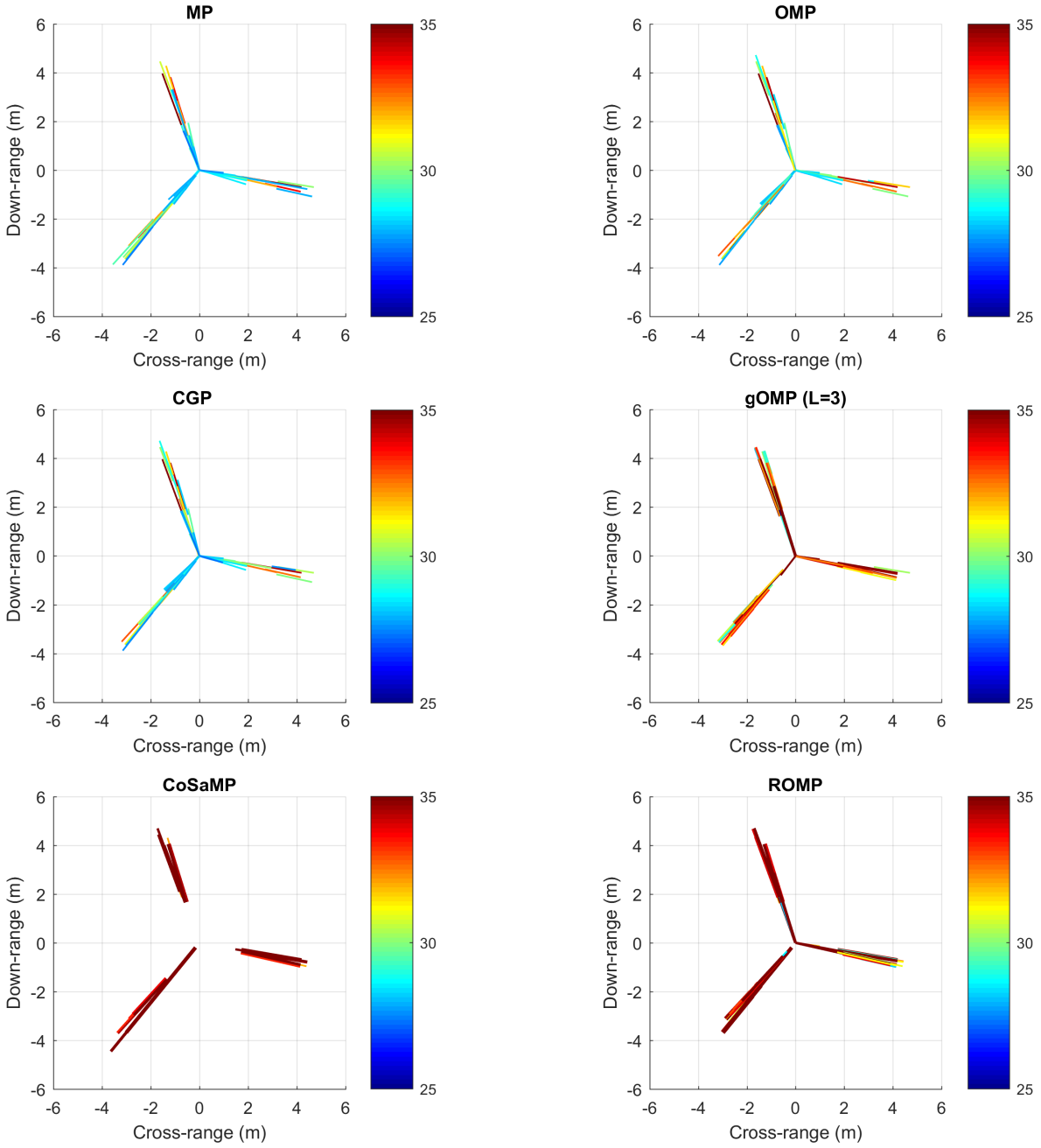


Figure 12: Scatter plots showing the tilted-wire atom parameters after processing real data with each of the greedy algorithms. Color bar indicates the amplitude of the atom coefficients.

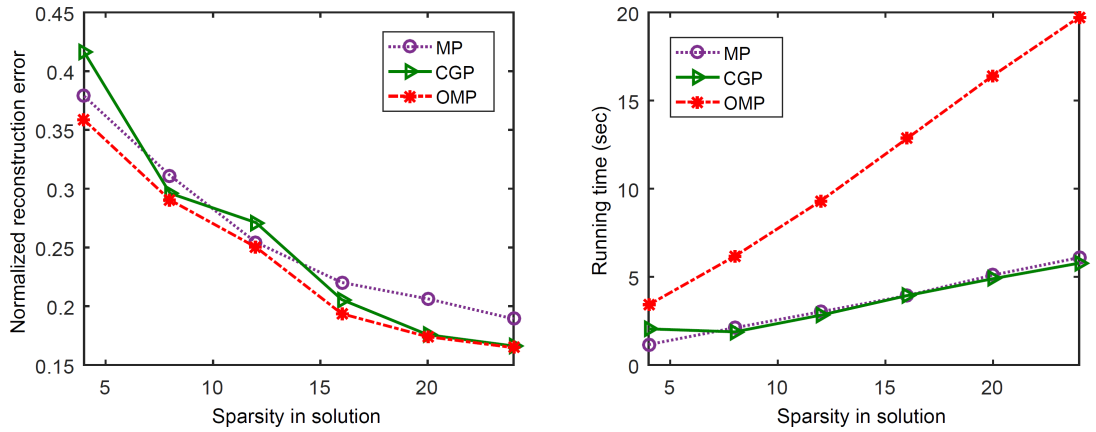


Figure 13: Normalised reconstruction error and running time after processing real data using MP, OMP and CGP.

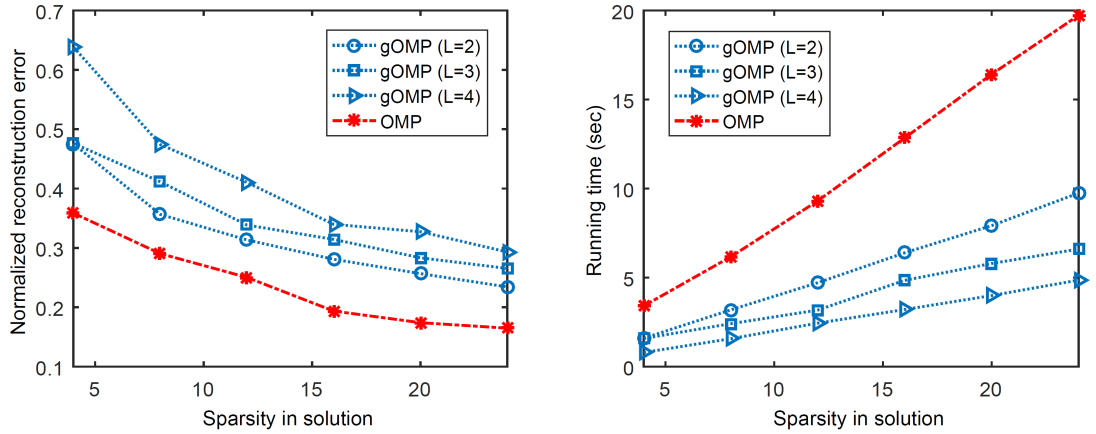


Figure 14: Normalised reconstruction error and running time after processing real data using gOMP.

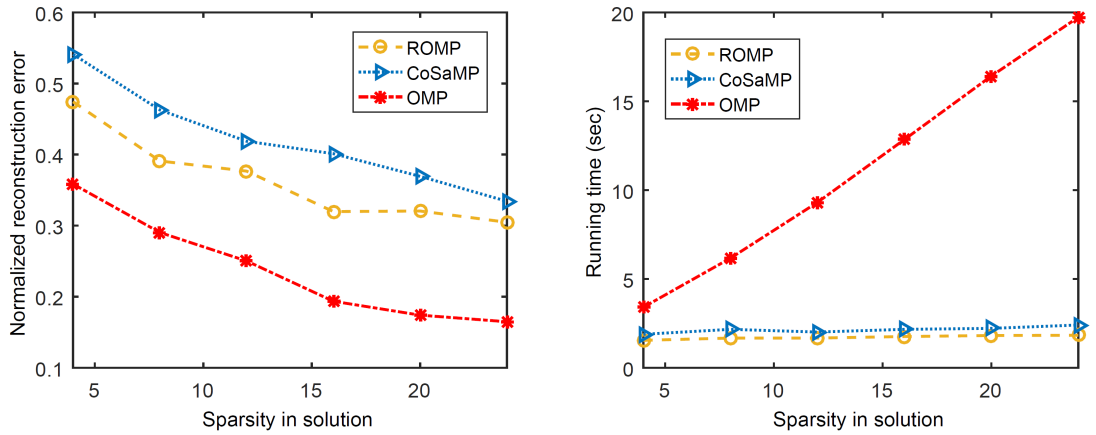


Figure 15: Normalised reconstruction error and running time after processing real data using ROMP, CoSaMP and OMP.

3.4 Further Discussion

The report documents a comprehensive survey of the currently known sparsity-based signal processing techniques. From the results for both simulated and real data, we observe that OMP, MP and CGP exhibit similar performance where OMP is slightly superior than MP and CGP. However, OMP takes much longer time to run compared to that of MP and CGP. Since CGP only performs a single gradient update at each iteration, CGP requires a sufficient number of iterations to converge to the 'correct' least-squares solution. For that reason, CGP performs worse than OMP for low sparsity levels of the solution vector. On the other hand, for sufficiently large sparsity in the solution vector, CGP performs very similar to OMP while it is more efficient in terms of computation.

The results also suggest that, by allowing the selection of multiple atoms in each iteration, gOMP gains significant reduction in runtime compared to OMP. However, this computational advantage comes at the expense of performance degradation as expected. A similar trend was observed for stOMP, SWOMP and SWCGP. The performance of stOMP, SWOMP and SWCGP, however, is not shown here as they are similar, at least for these examples. In addition, although providing very strong theoretical guarantees, it is observed that ROMP, SP and CoSaMP do not perform well compared to OMP. Note that, due to the similarity with SP and CoSaMP, only the performance of CoSaMP is included in this report. The common feature of gOMP, SWOMP, SWCGP, ROMP, SP and CoSaMP is that they select multiple new atoms at a time. However, for the particular problem of radar imaging of rotating blades, selecting multiple atoms is not desirable because the function dictionary $\mathbf{G}$ may be constructed from fairly dense grids of L, r, ψ and α. As the result, along with the correct atom, these algorithms may select other atoms which are closely located with the correct one, disrupting the accuracy of the representation.

Figures 16-21 showing the evolution of collected atoms with iteration index provide further insight into the behaviour of the techniques. Here, the algorithms for MP, OMP, CGP, gOMP and ROMP terminate when 10 atoms are selected while CoSaMP terminates when the l_2-norm of the signal reconstruction error dips below a certain threshold. Note that the stopping criterion based on the sparsity level in the solution vector is not appropriate for CoSaMP because it maintains a fixed sparsity level in the solution vector over iterations by removing insignificant atoms via the pruning step. In Figures 16-21, the magnitude of coefficient is color coded and the coordinate system has been rotated so that the blade is oriented upward. As shown in Figures 16-18, for the greedy algorithms (i.e., MP, OMP and CGP) which select a single atoms at a time, the coefficients of the collected atoms are very stable over iterations. On the other hand, gOMP, ROMP and CoSaMP select a group of atoms which are closely located as shown in Figures 19-21. As a result, along with the 'most matched' atom, these algorithms tend to also select other incorrect atoms 'nearby'. This is a plausible explanation for the performance degradation of gOMP, ROMP and CoSaMP compared to OMP.

We have also carried out other simulation runs where the true scatterers are more separated, the performances of the techniques become more comparable; however, this does deviate from the current interest and will be further investigated in future studies.

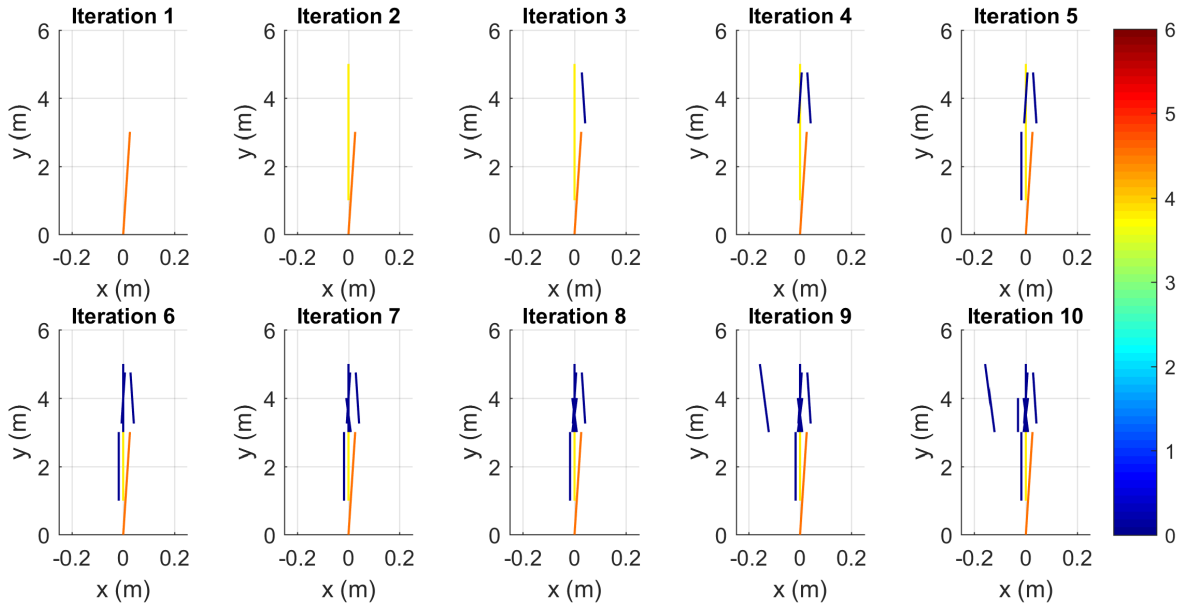


Figure 16: Scatter plots showing the evolution of atom parameters extracted using MP.

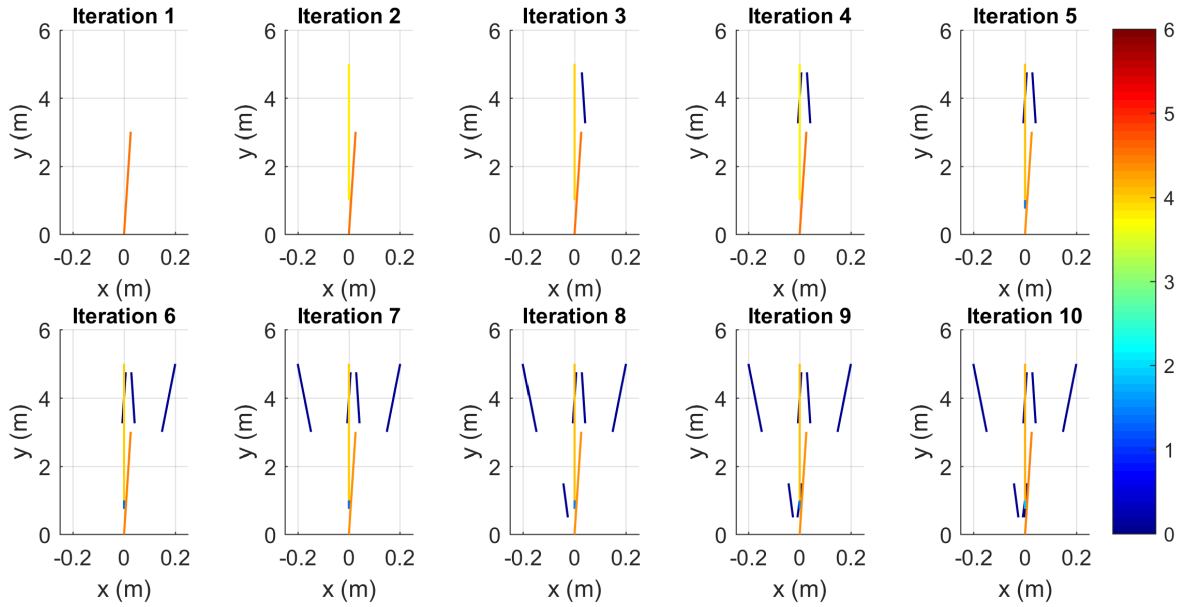


Figure 17: Scatter plots showing the evolution of atom parameters extracted using OMP.

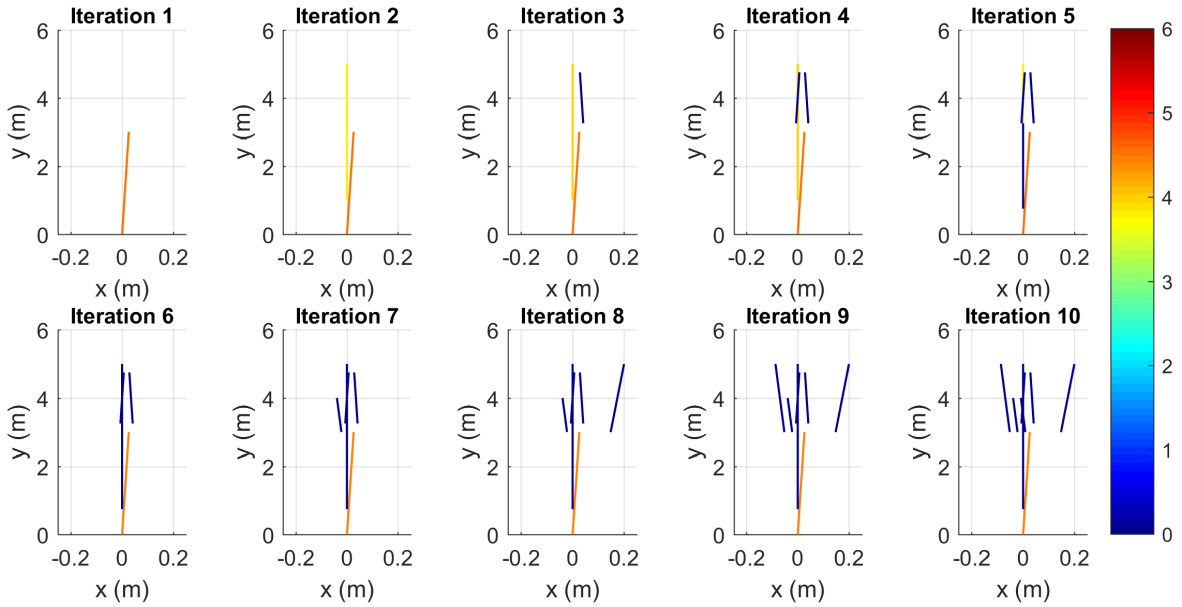


Figure 18: Scatter plots showing the evolution of atom parameters extracted using CGP.

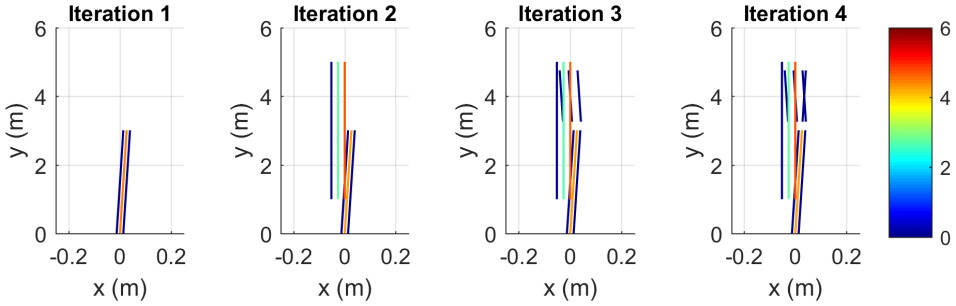


Figure 19: Scatter plots showing the evolution of atom parameters extracted using gOMP.

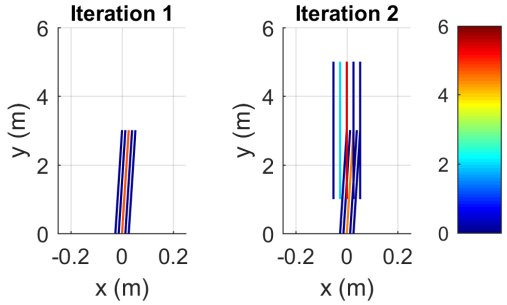


Figure 20: Scatter plots showing the evolution of atom parameters extracted using ROMP.

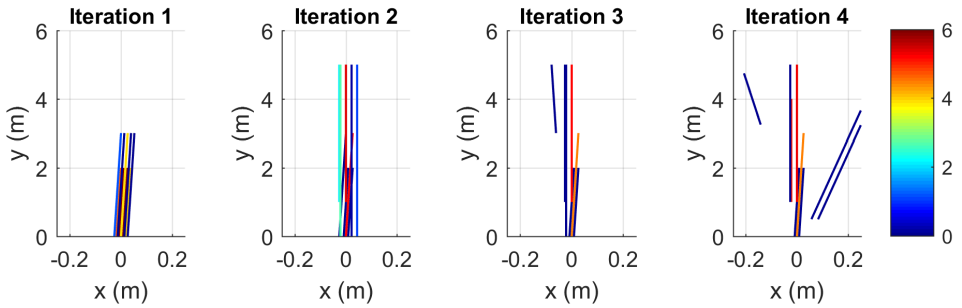


Figure 21: Scatter plots showing the evolution of atom parameters extracted using CoSaMP.

4 Conclusion

This report is part of a study on techniques for the analysis of radar backscatter from fast rotating blade-like objects using sparsity-based approaches in which the function dictionary is based on the exact solution of the tilted-wire scatterer model in two spatial dimensions. Three classes of sparse recovery algorithms were reviewed, based on the greedy, l_1-minimisation and l_p-minimisation ($p < 1$) methods. However, most of the report focuses on the performance and computational cost of the greedy algorithms.

Using both simulated and real radar data, this preliminary study found that those greedy algorithms that select multiple atoms at each iteration perform poorly in terms of high error in the signal reconstruction, although the computational cost is generally many times lower; these include the techniques of gOMP, SWOMP, SWCGP, ROMP, SP and CoSaMP. On the contrary, the techniques of MP, OMP, and CGP which select only one atom at each iteration perform better in terms of representation error, although with a higher computational cost. In particular, the CGP is computationally much more efficient than the OMP and yet exhibits a very similar performance to the OMP. Thus, in terms of performance-versus-computational cost trade-off, the CGP is the algorithm of choice for sparse analysis of narrowband radar backscatter from helicopter blades. This conclusion is valid at least for the cases where high correlation exists among the atoms of the underlying dictionary, which arises from closely spaced scatterers in the spatial domain.

Further investigation is currently underway and will be reported in future publications.

Acknowledgements

This work was successfully undertaken under Research Agreement (ID: 6360) between DST Group and University of South Australia. The authors would like to especially thank Dr Dmitri Kamenetsky of the Intelligence Analytics Branch, NSID, for his very detailed, critical, and helpful review, and also Dr Leigh Powis and Dr Andrew Shaw (Research Leader Surveillance and Reconnaissance Systems, NSID) for their review, support and leadership.

5 References

1. Cetin, M. and Karl, W. C. (Apr. 2001) 'Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization'. In: *IEEE Trans. Image Process.* **10** (4), 623–631.
2. Cetin, M., Karl, W. C. and Willsky, A. S. (2006) 'Feature-preserving regularization method for complex-valued inverse problems with application to coherent imaging'. In: *Optical Engineering* **45** (1), 017003 1–017003 11.
3. Cetin, M., Karl, W. C. and Castanon, D. A. (Oct. 2003) 'Feature enhancement and ATR performance using nonquadratic optimization-based SAR imaging'. In: *IEEE Trans. Aerosp. Electron. Syst.* **39** (4), 1375–1395.
4. Moses, R. L., Potter, L. C. and Cetin, M. (Apr. 2004) 'Wide-angle SAR imaging'. In: *Proc. SPIE Algorithms Synthetic Aperture Radar Imagery XI*. Vol. 5427, 164–175.
5. Cetin, M. and Moses, R. L. (Mar. 2005) 'SAR imaging from partial-aperture data with frequency-band omissions'. In: *Proc. SPIE Algorithms Synthetic Aperture Radar Imagery XII*. Vol. 5808, 32–43.
6. Stojanovic, I., Cetin, M. and Karl, W. C. (2008) 'Joint space aspect reconstruction of wide-angle SAR exploiting sparsity'. In: *Proc. SPIE Defense Security Symp. Algorithms Synthetic Aperture Radar Imagery XV*. Vol. 6970, 697005 1–697005 12.
7. Varshney, K. R. et al. (Aug. 2008) 'Sparse representation in structured dictionaries with application to synthetic aperture radar'. In: *IEEE Trans. Signal Process.* **56** (8), 3548–3561.
8. Ramakrishnan, N., Ertin, E. and Moses, R. L. (Apr. 2007) 'Joint enhancement of multichannel SAR data'. In: *Proc. SPIE Algorithms Synthetic Aperture Radar Imagery XIV*. Vol. 6568, 65680B 1–65680B 9.
9. Ramakrishnan, N., Ertin, E. and Moses, R. L. (Aug. 2010) 'Enhancement of coupled multichannel images using sparsity constraints'. In: *IEEE Trans. Image Process.* **19** (8), 2115–2126.
10. Himed, B. et al. (2001) 'Tomography of moving targets (TMT)'. In: *Proc. SPIE Sensors, Syst., Next-Gen. Satellites V*. Vol. 4540. 1, 608–619.
11. Stojanovic, I. and Karl, W. C. (Feb. 2010) 'Imaging of moving targets with multi-static SAR using an overcomplete dictionary'. In: *IEEE J. Sel. Topics Signal Process.* **4** (1), 164–176.
12. Chen, V. C. et al. (Jan. 2006) 'Micro-Doppler effect in radar: phenomenon, model, and simulation study'. In: *IEEE Trans. Aerosp. Electron. Syst.* **42** (1), 2–21.
13. Tran, H. T. et al. (Sept. 2013) 'Microwave radar imaging of rotating blades'. In: *Proc. Int. Conf. Radar*, 202–207.
14. Melino, R., Kodituwakku, S. and Tran, H. T. (Oct. 2015) 'Orthogonal matching pursuit and matched filter techniques for the imaging of rotating blades'. In: *Proc. IEEE Radar Conf.* 1–6.

15. Natarajan, B. K. (1995) 'Sparse approximate solutions to linear systems'. In: *SIAM Journal on Computing* **24** (2), 227–234.

16. Donoho, D. L. (Apr. 2006) 'Compressed sensing'. In: *IEEE Trans. Inf. Theory* **52** (4), 1289–1306.

17. Candes, E. J. and Tao, T. (Dec. 2005) 'Decoding by linear programming'. In: *IEEE Trans. Inf. Theory* **51** (12), 4203–4215.

18. Candes, E. J. (2008) 'The restricted isometry property and its implications for compressed sensing'. In: *Comptes Rendus Mathematique* **346** (910), 589–592.

19. Candes, E. J., Romberg, J. and Tao, T. (Feb. 2006) 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information'. In: *IEEE Trans. Inf. Theory* **52** (2), 489–509.

20. Tropp, J. A. (Mar. 2006) 'Just relax: convex programming methods for identifying sparse signals in noise'. In: *IEEE Trans. Inf. Theory* **52** (3), 1030–1051.

21. Y. C. Eldar, G. K., ed. (2012) *Compressed Sensing: Theory and Applications*. Cambridge University Press. New York.

22. Potter, L. C. et al. (June 2010) 'Sparsity and compressed sensing in radar imaging'. In: *Proc. IEEE* **98** (6), 1006–1020.

23. Needell, D. and Vershyni, R. (2009) 'Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit'. In: *Foundations of Computational Mathematics* **9** (3), 317–334.

24. Dai, W. and Milenkovic, O. (May 2009) 'Subspace pursuit for compressive sensing signal reconstruction'. In: *IEEE Trans. Inf. Theory* **55** (5), 2230–2249.

25. Donoho, D. L. et al. (Feb. 2012) 'Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit'. In: *IEEE Trans. Inf. Theory* **58** (2), 1094–1121.

26. Wang, J., Kwon, S. and Shim, B. (Dec. 2012) 'Generalized orthogonal matching pursuit'. In: *IEEE Trans. Signal Process.* **60** (12), 6202–6216.

27. Mallat, S. G. and Zhang, Z. (Dec. 1993) 'Matching pursuits with time-frequency dictionaries'. In: *IEEE Trans. Signal Process.* **41** (12), 3397–3415.

28. Pati, Y. C., Rezaiifar, R. and Krishnaprasad, P. S. (Nov. 1993) 'Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition'. In: *Proc. 27th Asilomar Conf. Signals, Syst., Comput.* Vol. 1, 40–44.

29. Blumensath, T. and Davies, M. E. (Nov. 2009) 'Stagewise weak gradient pursuits'. In: *IEEE Trans. Signal Process.* **57** (11), 4333–4346.

30. Blumensath, T. and Davies, M. E. (June 2008) 'Gradient pursuits'. In: *IEEE Trans. Signal Process.* **56** (6), 2370–2382.

31. Needell, D. and Vershynin, R. (Apr. 2010) 'Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit'. In: *IEEE J. Sel. Topics Signal Process.* **4** (2), 310–316.

32. Needell, D. and Tropp, J. A. (2009) 'CoSaMP: iterative signal recovery from incomplete and inaccurate samples'. In: *Appl. Comp. Harmonic Anal.* **26** (3), 301–321. ISSN: 1063-5203.

33. Chartrand, R. (Oct. 2007) 'Exact reconstruction of sparse signals via nonconvex minimization'. In: *IEEE Signal Process. Lett.* **14** (10), 707–710.

34. Chartrand, R. and Staneva, V. (2008) 'Restricted isometry properties and nonconvex compressive sensing'. In: *Inverse Problems* **24** (3), 1–14.

35. Chartrand, R. and Yin, W. (Mar. 2008) 'Iteratively reweighted algorithms for compressive sensing'. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.* Las Vegas, USA, 3869–3872.

36. Donoho, D. L. et al. (Feb. 2012) 'Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit'. In: *IEEE Trans. Inf. Theory* **58** (2), 1094–1121.

37. Tropp, J. A. and Gilbert, A. C. (Dec. 2007) 'Signal recovery from random measurements via orthogonal matching pursuit'. In: *IEEE Trans. Inf. Theory* **53** (12), 4655–4666.

38. Rauhut, H. (May 2008) 'On the impossibility of uniform sparse reconstruction using greedy methods'. In: *Sampl. Theory Signal Image Process.* **7** (2), 197–215.

39. Donoho, D. L. (June 2006) 'For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution'. In: *Commun. Pure Appl. Math.* **59** (6), 797–829.

40. Tropp, J. A. (Oct. 2004) 'Greed is good: algorithmic results for sparse approximation'. In: *IEEE Trans. Inf. Theory* **50** (10), 2231–2242.

41. Donoho, D. L. and Huo, X. (Nov. 2001) 'Uncertainty principles and ideal atomic decomposition'. In: *IEEE Trans. Inf. Theory* **47** (7), 2845–2862.

42. Cai, T. T. and Wang, L. (July 2011) 'Orthogonal matching pursuit for sparse signal recovery with noise'. In: *IEEE Trans. Inf. Theory* **57** (7), 4680–4688.

43. Baraniuk, R. et al. (2008) 'A simple proof of the restricted isometry property for random matrices'. In: *Constructive Approximation* **28** (3), 253–263.

44. Bandeira, A. S. et al. (June 2013) 'Certifying the restricted isometry property is hard'. In: *IEEE Trans. Inf. Theory* **59** (6), 3448–3450.

45. Davenport, M. A. and Wakin, M. B. (Sept. 2010) 'Analysis of orthogonal matching pursuit using the restricted isometry property'. In: *IEEE Trans. Inf. Theory* **56** (9), 4395–4401.

46. Wang, J. and Shim, B. (Sept. 2012) 'On the recovery limit of sparse signals using orthogonal matching pursuit'. In: *IEEE Trans. Signal Process.* **60** (9), 4973–4976.

47. Golub, G. H. and Loan, C. F. V. (1996) *Matrix Computations*. Johns Hopkins University Press. London.

48. Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*. Cambridge University Press. Cambridge, UK.

49. Nocedal, J. and Wright, S. (2006) *Numerical Optimization*. Springer Science & Business Media. New York.

50. Figueiredo, M. A. T., Nowak, R. D. and Wright, S. J. (Dec. 2007) 'Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems'. In: *IEEE J. Sel. Topics Signal Process.* **1** (4), 586–597.

51. Tropp, J. A. (2006) 'Algorithms for simultaneous sparse approximation. Part II: Convex relaxation'. In: *Signal Process.* **86** (3). Sparse Approximations in Signal and Image ProcessingSparse Approximations in Signal and Image Processing, 589–602.

52. Tibshirani, R. (1996) 'Regression shrinkage and selection via the LASSO'. In: *J. R. Statist. Soc. B,* 267–288.

53. Chen, S. S., Donoho, D. L. and Saunders, M. A. (2001) 'Atomic decomposition by basis pursuit'. In: *SIAM Review* **43** (1), 129–159.

54. Rockafellar, R. T. (1970) *Convex Analysis*. Princeton University Press. New Jersey.

55. Galatsanos, N. P. and Katsaggelos, A. K. (July 1992) 'Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation'. In: *IEEE Trans. Image Process.* **1** (3), 322–336.

56. Golub, G. H., Heath, M. and Wahba, G. (1979) 'Generalized cross-validation as a method for choosing a good ridge parameter'. In: *Technometrics* **21** (2), 215–223.

57. Rao, B. D. and Kreutz-Delgado, K. (Jan. 1999) 'An affine scaling methodology for best basis selection'. In: *IEEE Trans. Signal Process.* **47** (1), 187–200.

58. Yoon, S.-H., Kim, B. and Kim, Y.-S. (Oct. 2000) 'Helicopter classification using time-frequency analysis'. In: *Electronics Letters* **36** (22), 1871–1872.

59. Gravityslave *website: www.youtube.com.au/watch?v=Ug6W7 tafnc.*

DEFENCE SCIENCE AND TECHNOLOGY GROUP DOCUMENT CONTROL DATA			1. DLM/CAVEAT (OF DOCUMENT)
2. TITLE A Review of Sparsity-Based Methods for Analysing Radar Returns from Helicopter Rotor Blades		3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)	
4. AUTHORS Ngoc Hung Nguyen, Hai-Tan Tran, Kutluyıl Doğançay and Rocco Melino		5. CORPORATE AUTHOR Defence Science and Technology Group PO Box 1500 Edinburgh, South Australia 5111, Australia	
6a. DST Group NUMBER DST-Group–TR–3292	6b. AR NUMBER 016-687	6c. TYPE OF REPORT Technical Report	7. DOCUMENT DATE September, 2016
8. Objective ID	9. TASK NUMBER	10. TASK SPONSOR	
13. DST Group Publications Repository `http://dspace.dsto.defence.gov.au/dspace/`		14. RELEASE AUTHORITY Chief, National Security and ISR Division	
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT *Approved for Public Release* OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111			
16. DELIBERATE ANNOUNCEMENT			
17. CITATION IN OTHER DOCUMENTS No Limitations			
18. RESEARCH LIBRARY THESAURUS Sparse Representation, Matching Pursuit, Orthogonal Matching Pursuit, Conjugate Gradient Pursuit, Helicopter Rotor Blades, Tilted-wire Model, Radar Imaging			

19. ABSTRACT

Radar imaging of rotating blade-like objects, such as helicopter rotors, using narrowband radar has lately been of significant interest; these objects cannot be adequately described by the classic point-scatterer model. Recently, a novel 'tilted-wire' scatterer model has been developed that can provide an accurate and sparse representation of radar returns from such objects.

Following a literature review on compressed sensing algorithms, covering both greedy and l_p minimisation methods ($0 < p \leq 1$), the report focuses on a comparative study of various greedy pursuit algorithms, using both simulated and real radar data, with a particular emphasis on the use of the tilted-wire scatterer model. It is observed that the greedy algorithms that select multiple atoms at the matched-filtering stage do not perform well when the atoms used in the dictionary are significantly correlated. Amongst the greedy algorithms, Orthogonal Matching Pursuit (OMP) exhibits the best performance, closely followed by Conjugate Gradient Pursuit (CGP), which has a much smaller computational complexity than OMP. In applications where the tilted-wire model requires large dictionaries and large CPI atoms, CGP is the preferred option.